



22

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/12, 15/62, C07K 14/47, 16/18, G01N 33/53</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 95/27059</b> <b>(43) International Publication Date:</b> 12 October 1995 (12.10.95)
<b>(21) International Application Number:</b> PCT/US95/04075 <b>(22) International Filing Date:</b> 31 March 1995 (31.03.95) <b>(30) Priority Data:</b> 08/222,619 - 31 March 1994 (31.03.94) US <b>(71) Applicants:</b> AMGEN INC. [US/US]; Amgen Center, 1840 Dehavilland Drive, Thousand Oaks, CA 91320-1789 (US). THE ROCKEFELLER UNIVERSITY [US/US]; 1230 York Avenue, New York, NY 10021-6399 (US). <b>(72) Inventors:</b> LICHENSTEIN, Henri, Stephen; 9586 Lucerne Street, Ventura, CA 93004 (US). LYONS, David, Edwin; 2027 Truett Circle, Thousand Oaks, CA 91320-1789 (US). WURFEL, Mark, Matsuo; 420 E. 70 Street, New York, NY 10021 (US). WRIGHT, Samuel, Donald; 2 Briar Close, Larchmont, NY 10538 (US). <b>(74) Agents:</b> ODRE, Steven, M. et al.; Amgen Inc., Amgen Center, 1840 Dehavilland Drive, Thousand Oaks, CA 91320-1789 (US).		<b>(81) Designated States:</b> AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> AFAMIN: A HUMAN SERUM ALBUMIN-LIKE PROTEIN  <b>(57) Abstract</b>  The invention relates to a novel human serum protein referred to as AFM, which has one or more activities in common with human serum albumin, human $\alpha$ -fetoprotein, or human vitamin D binding protein and which has an apparent molecular weight by SDS-PAGE of 87 kd; variants thereof; and related genes, vectors, cells and methods.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## AFAMIN: A HUMAN SERUM ALBUMIN-LIKE PROTEIN

Field of the Invention

5

Generally, the invention relates to the field of human serum proteins that are functionally and structurally similar to the related proteins: human serum albumin (ALB), human  $\alpha$ -fetoprotein (AFP), and vitamin D-binding protein (VDB).

10

Background of the Invention

The human serum proteins albumin (ALB),  $\alpha$ -feto-protein (AFP) and vitamin D binding protein (VDB) are known to be members of a multigene ALB family. All three proteins are found in serum where they mediate the transport of a wide variety of ligands. ALB binds fatty acids, amino acids, steroids, glutathione, metals, bilirubin, lysolecithin, hematin, prostaglandins and pharmaceuticals (for review, see 1). AFP binds fatty acids, bilirubin and metals (2, 3). VDB binds vitamin D and its metabolites as well as fatty acids, actin, C5a and C5a des Arg (4-7).

20

In addition to their transport capabilities, ALB family proteins possess a wide assortment of other functional activities. ALB is the main contributor to the colloid oncotic pressure of plasma, acts as a scavenger of oxygen-free radicals and can inhibit copper-stimulated lipid peroxidation, hydrogen peroxide release, and neutrophil spreading (1, 8-10). AFP has been implicated in the regulation of immune processes (11-14) and VDB can act as a co-chemotactic factor for neutrophils (6, 15) and as an activating factor for macrophages (16).

30

35

The serum levels of ALB family proteins are also known to be responsive to various pathological conditions. ALB is a negative acute phase protein (17) whose levels decrease in times of stress. AFP levels are elevated in women carrying fetuses with certain developmental disorders (18, 19) and in individuals with hepatocarcinoma, teratocarcinoma, hereditary tyrosinemia or ataxia-telangiectasia (20-24). VDB levels are decreased in patients with septic shock (25) or fulminant hepatic necrosis (26, 27).

ALB family members also have significant structural similarities. Homology has been observed at the primary amino acid sequence level and there is also a well-conserved pattern of Cys residues which predicts similar secondary structures (28-32). ALB family genes have similar exon/intron organizations (33-36) and all have been mapped to human chromosome 4 within the region 4q11-q22 (37, 38).

Human "Afamin" (abbreviated as "AFM") is a novel serum protein with a molecular weight of 87000 daltons. It shares strong similarity to albumin family members and has the characteristic pattern of disulfide bonds observed in this family. In addition, the gene maps to chromosome 4 as do other members of the albumin gene family. Thus, AFM is the fourth member of the albumin family of proteins. AFM cDNA was stably transfected into Chinese hamster ovary cells and recombinant protein (rAFM) was purified from conditioned medium. Both rAFM and AFM purified from human serum react with a polyclonal antibody that was raised against a synthetic peptide derived from the deduced amino acid sequence of AFM. It is expected that AFM will have properties and biological activities in common with ALB, AFP, and VDB.

Publications relating to "Background of the Invention"

1. Peters, T. Jr., *Adv. Protein Chem.* **37**, 161 - 245 (1985).
- 5 2. Parmelee, D.C., Evenson, M.A., and Deutsch, H.F., *J. Biol. Chem.* **253**, 2114 - 2119 (1978).
3. Berde, C.B., Nagai, M., Deutsch, H.F., *J. Biol.*  
10 *Chem.* **254**, 12609 - 12614 (1979).
4. Daiger, S.P., Schanfield, M.S., and Cavalli-Sforza, L.L., *Proc. Nat. Acad. Sci. U.S.A.* **72**, 2076 - 2080 (1975).
- 15 5. Van Baelen, H., Bouillon, R., and De Moor, P., *J. Biol. Chem.* **255**, 2270 - 2272 (1980).
6. Kew, R.R., and Webster, R.O., *J. Clin. Invest.* **82**,  
20 364 - 369 (1988).
7. Williams, M.H., Van Alstyne, E.L., and Galbraith, R.M., *Biochem. Biophys. Res. Commun.* **153**, 1019 - 1024 (1988).
- 25 8. Holt, M.E., Ryall, M.E.T., and Campbell, A.K., *Br. J. exp. Path.* **65**, 231 - 241 (1948).
9. Gutteridge, J.M.C., *Biochim. Biophys ACTA* **869**, 119  
30 -127 (1986).
10. Nathan, C., Xie, Q-W., Halbwachs-Mecarelli, L. and Jin, W.W., *J. Cell Biol.* **122**, 243-256 (1993).
- 35 11. Yachnin, S., *Proc. Natl. Acad. Sci. U.S.A.* **73**, 2857 - 2861 (1976).

12. Auer, I.O., and Kress, H.G., *Cell. Immunol.* **30**, 173 - 179 (1977).
- 5 13. Alpert, E., Dienstag, J.L., Sepersky, S., Littman, B., and Rocklin, R., *Immunol. Commun.* **7**, 163 - 185 (1978).
- 10 14. Chakraborty, M., and Mandal, C., *Immunol. Invest.* **22**, 329 - 339 (1993).
- 15 15. Perez, H.D., Kelly, E., Chenoweth, D., and Elfman, F., *J. Clin. Invest.* **82**, 360 - 363 (1988).
- 16 16. Yamamoto, N. and Homma, S., *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8539 - 8543 (1991).
- 20 17. Koj, A., in *The Acute Phase Response to Injury and Infection 1* (Gordon, A.H., and Koj, A. eds) pp. 45 - 160 (1985).
18. Brock, D.J.H., and Sutcliffe, R.G., *Lancet* **2**, 197 - 199 (1972).
- 25 19. Allan, L.D., Ferguson-Smith, M.A., Donald, I., Sweet, E.M. and Gibson, A.A.M., *Lancet* **2** 522 - 525 (1973).
- 30 20. Waldmann, T.A., and McIntire, K.R., *Lancet* **2**, 1112 - 1115 (1972).
- 35 21. Belanger, L., Belanger, M., Prive, L., Larochelle, J., Tremblay, M., and Aubin, G., *Pathol. Biol.* **21**, 449 - 455 (1973).
22. Belanger, L., *Pathol. Biol.* **21**, 457 - 463 (1973).

23. Ruoslahti, E., Pihto, H., and Seppala, M.,  
*Transplant. Rev.* **20**, 38 - 60 (1974).
- 5 24. Tamaoki, T., and Fausto, in *Recombinant DNA and  
Cell Proliferation* (Stein G. and Stein J. eds.) pp. 145  
- 168 (1984).
25. Lee, W.M., Reines, D., Watt, G.H., Cook, J.A.,  
10 Wise, W.C., Halushka, P.V., and Galbraith, R.M., *Circ.  
Shock.* **28**, 249 - 255 (1989).
26. Young, W.O., Goldschmidt-Clermont, P.J., Emerson,  
D.L., Lee, W.M., Jollow, D.J., and Galbraith, R.M., *Lab.*  
15 *Clin. Med.* **110**, 83 - 90 (1987).
27. Goldschmidt-Clermont, P.J., Lee, W.M., and  
Galbraith, R.M., *Gastroenterology* **94**, 1454 - 1458  
(1988).
- 20 28. Lawn, R.M., Adelman, J., Bock, S.C., Franke, A.E.,  
Houck, C.M., Najarian, R.C., Seeburg, P.H., and Wion,  
K.L., *Nucleic Acid Res.* **9**, 6103 - 6114 (1981).
- 25 29. Dugaiczky, A., Law, S.W., and Dennison, O.E., *Proc.  
Natl. Acad. Sci. U.S.A.* **79**, 71 - 75 (1982).
- 30 30. Morinaga, T., Sakai, M., Wegmann, T.G., and  
Tamaoki, T., *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4604 -  
4608 (1983).
31. Cooke, N.E. and David, E.V., *J. Clin. Invest.* **76**,  
2420 - 2424 (1985).

32. Yang, F., Brune, J.L., Naylor, S.L., Cupples, R.L., Naberhaus, K.H., and Bowman, B.H., *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7994 - 7998 (1985).
- 5 33. Sakai, M., Morinaga, T., Urano, Y., Watanabe, K., Wegmann, T.G., and Tamaoki, T., *J. Biol. Chem.* **260**, 5055 - 5060 (1985).
- 10 34. Minghetti, P.P., Ruffner, D.E., Kuang, W-J., Dennison, O.E., Hawkins, J.W., Beattie, W.G., and Dugaiczky, A., *J. Biol. Chem.* **261**, 6747 - 6757 (1986).
- 15 35. Gibbs, P.E.M., Zielinski, R., Boyd, C., and Dugaiczky, A., *Biochemistry* **26**, 1332 - 1343 (1987).
36. Witke, W.F., Gibbs, P.E.M., Zielinski, R., Yang, F., Bowman, B.H., and Dugaiczky, A., *Genomics* **16**, 751 - 754 (1993).
- 20 37. Mikkelsen, M., Jacobsen, P. and Henningsen, K., *Hum. Hered.* **27**, 105 - 107 (1977).
38. Harper, M.E., and Dugaiczky, A., *Am. J. Hum. Genet.* **35**, 565-572 (1983).

25

The sections below contain a summary of background information that is currently available on ALB, AFP, and VDB and contains lists of additional publications relating to these known proteins.

30

#### I. Human Serum Albumin

Human serum albumin is an important factor in the regulation of plasma volume and tissue fluid balance through its contribution to the colloid osmotic pressure of plasma. Albumin normally constitutes 50-60% of

35



plasma proteins and because of its relatively low molecular weight (66,300-69,000), exerts 80-85% of the colloidal osmotic pressure of the blood.

The best known functions of ALB involve  
5 regulation of transvascular fluid flux and hence, intra  
and extravascular fluid volumes and transport of lipid  
and lipid-soluble substances. ALB solutions are  
frequently used for plasma volume expansion and  
maintenance of cardiac output in the treatment of  
10 certain types of shock or impending shock including  
those resulting from burns, surgery, hemorrhage, or  
other trauma or conditions in which a circulatory volume  
deficit is present. Transfusions of whole blood or red  
blood cells also may be necessary, depending on the  
15 severity of red blood cell loss.

Intravenous (IV) administration of  
concentrated ALB solutions causes a shift of fluid from  
the interstitial spaces into the circulation and a  
slight increase in the concentration of plasma proteins.  
20 When administered IV to a well-hydrated patient, each  
volume of 25% ALB solution draws about 3.5 volumes of  
additional fluid into the circulation within 15 minutes,  
reducing hemoconcentration and blood viscosity. In  
patients with reduced circulating blood volumes (as from  
25 hemorrhage or loss of fluid through exudates or into  
extravascular spaces), hemodilution persists for many  
hours, but in patients with normal blood volume, excess  
fluid and protein are lost from the circulation within a  
few hours. In dehydrated patients, ALB generally  
30 produces little or no clinical improvement unless  
additional fluids are administered.

Although ALB contains some bound amino acids,  
it provides only modest nutritive effect. ALB binds and  
functions as a carrier of intermediate metabolites  
35 (including bilirubin), trace metals, some drugs, dyes,  
fatty acids, hormones, and enzymes, thus affecting the

transport, inactivation, and/or exchange of tissue products.

ALB is also involved in a number of other vital functions, some of which have only recently been suggested and perhaps others which are as yet unrecognized. Among recognized unique features of albumin are: a) binding, and hence, inactivation of toxic products; b) regulation of the plasma and interstitial fluid concentrations of endogenous and exogenously administered substances and drugs; c) involvement in anticoagulation; d) maintenance of microvascular permeability to protein; and e) scavenging of free radicals and prevention of lipid peroxidation. This latter property may prove to be critically important, particularly in inflammatory disease states in which free radicals are thought to be a major culprit in direct damage due to tissue oxidation, and indirect tissue damage due to inactivation of important antiproteases such as  $\alpha_1$ -PI and AT-III.

The following is a more detailed summary of the many uses for ALB that have been reported in the literature:

A. Functions of ALB

- Contributes to colloid osmotic pressure and thus prevents water loss from circulation;
- Aids in transport, distribution, metabolism of fatty acids (primarily long chain), amino acids (Cys and Trp), steroids, glutathione, metals (Ca, Zn), bilirubin, lysolecithin, hematin, prostaglandins and pharmaceuticals to liver, intestine, kidney and brain presumably through specific albumin receptors that have been identified on the endothelium;

- Serves as a reservoir for fatty acids intra and extravascularly (60% of the ALB is found extravascularly);  
5
- Modification of doxorubicin (DXR) by conjugating it to bovine serum albumin (BSA) improved chemotherapeutic efficiency of DXR presumably by decreasing efflux of BSA-DXR compared to DXR alone (in  
10 animal models), suggesting a similar use with ALB;
- Inhibits Cu-stimulated lipid peroxidation and hemolysis of erythrocyte membranes (acts as antioxidant);  
15
- Scavenges HOCl and peroxy radicals;
- Prevents peroxidation of fatty acids by binding to them;  
20
- May exert a protective effect in body fluids that have little endogenous antioxidant protection (e.g., eye and cerebrospinal fluids);
- In urine, high levels of ALB are diagnostic for detection of early renal pathology in  
25 diabetics;
- Administered to combat shock and given to  
30 neonates with respiratory distress syndrome;
- Administered as a vehicle for hematin to treat acute intermittent porphyria;
- Used in tissue culture in place of whole  
35 serum;

- Enhances effectiveness of superoxide dismutase (SOD) when coupled to SOD through enhanced serum half-life;

5

- In microsphere form, ALB is useful as a carrier of therapeutic agents;

- Inhibits hydrogen peroxide release and neutrophil spreading.

10

#### B. Publications relating to ALB

15 38. Peters, Theodore in *ALBUMIN An Overview and Bibliography*, Second Edition, 1992.

39. American Hospital Formulary Service Drug Information, *Blood Derivatives*, 762-763 (1992).

20

40. Yamashita, T., et al., *Biochem. Biophys. Res. Commun.* **191** (2), 715-720 (1993).

41. Candlish, John K., *Pathology* **25**, 148-151 (1993).

25

42. Ohkawa, K., et al., *Cancer Research* **54**, 4238-4242 (1993).

43. He, Xiao Min and Carter, Daniel C., *Nature* **358**, 209-215 (1992).

30

44. Brown, J. M., et al., *Inflammation* **13** (5), 583-589 (1989).

35 45. Emerson, T. E., *Critical Care Medicine* **17** (7), 690-694 (1989).

46. Halliwell, Barry, *Biochem. Pharmacol.* **37** (4), 569-571 (1988).

5 47. Holt, M. E., et al., *Br. J. Exp. Path.* **65**, 231-241 (1984).

## II. Alpha Fetoprotein

10           Alpha-fetoprotein (AFP; molecular weight 70,000) is a major serum protein produced during development and is produced primarily by the fetal liver and yolk sac cells. Its synthesis decreases markedly after birth and only trace amounts are present in the  
15 serum of adults. Increased adult serum levels are a sign of hepatoma or yolk sac tumor, since these tumors produce AFP. The specific associations of AFP with fetal development as well as the above type of malignancies has attracted much interest and many  
20 studies have been done on the structure of AFP and its gene, the regulation of gene expression, and biological functions.

          Similar to ALB, AFP has been shown to bind various ligands such as unsaturated fatty acids,  
25 estrogens, bilirubin, copper and nickel ions, and others. AFP also has been claimed to regulate immune processes in a variety of systems from many different laboratories, although the results are controversial.

          The following is a more detailed list of uses  
30 for AFP that are available in the literature:

### A. Functions of AFP

          • Binds unsaturated fatty acids, estrogens,  
35 bilirubin, Cu, Ni;

- Elevated levels in amniotic fluid of pregnant women indicative of fetal malformations;
- High levels also found in hereditary tyrosinemia and ataxia-telangiectasia (autosomal recessive disorder characterized by a defect in tissue differentiation of thymus and liver);
- Inhibits NK cell activity;
- Induces T suppressor cells;
- Inhibits mitogenic responses of lymphocytes to PHA and ConA;
- Inhibits T cell proliferation to Ia determinants;
- Decreases macrophage phagocytosis and Ia expression;
- Inhibits FSH-mediated estradiol production by porcine granulosa cells;
- Enhances growth-factor mediated cell proliferation of porcine granulosa cells.

B. Publications relating to AFP

- 48. Suzuki, Y., et al., *J. Clin. Invest.* 90, 1530-1536 (1992).
- 49. Sakai, M., et al., *J. Biol. Chem.* 260 (8), 5055-5060 (1985).

50. EPO Patent Application No. 0353814,  
February 7, 1990.

### III. Vitamin-D Binding Protein

5

The group-specific component (Gc; VDB) is an  $\alpha_2$ -globulin of molecular weight 51,000. It is synthesized in the liver and is the major vitamin D-binding protein in plasma. VDB appears in human  
10 populations as three common genetic phenotypes: Gc1, Gc2, and Gc2-1. VDB has also been reported to bind G-actin and to be spatially associated with IgG on lymphocyte membranes.

The following is a more detailed list of uses  
15 for VDB that are available in the literature:

#### A. Functions of VDB

- Binds seco-steroid, vitamin D and the  
20 derivatives 25-hydroxy vitamin D and 1,25 hydroxy vitamin D, possibly for transport in plasma;
- 1,25 vitamin D can differentiate  
monocytes and VDB prevents this;  
25
- Binds actin (prevents assembly of actin  
polymers);
- Binds unsaturated fatty acids (e.g.,  
30 arachidonic acid);
- Binds C5a and C5a des Arg to act as a  
cochemotactic factor for neutrophils;
- 35 • Acts as an activating factor for  
macrophages.

B. Publication relating to VDB

51. Watt, G. H., et al., *Circulatory Shock*  
5 28, 279-291 (1989).

The protein of the present invention, AFM, bears a strong similarity in structure to ALB, AFP, and VDB, and is therefore expected to share the above  
10 utilities and activities with the known proteins discussed above.

Summary of the Invention

15 In the course of experiments designed to purify a serum protein which could inhibit the binding of lipopolysaccharide (LPS)-coated erythrocytes to human macrophages, the inventors purified a novel human protein that co-purifies with apolipoprotein A1 (ApoA1).  
20 The novel protein has an apparent molecular weight of 87,000 when run on SDS-PAGE and is designated as AFM. Herein, the inventors describe the cloning of the cDNA for AFM and demonstrate that AFM has a striking similarity, both structurally and functionally, to other  
25 members of the ALB family. In addition, the inventors purified AFM from the serum-free conditioned medium of CHO D<sup>-</sup> cells transfected with the cDNA for AFM, thus allowing the study of AFM in the absence of ApoA1.

Based on the above, the present invention  
30 provides purified and isolated polynucleotides (e.g., DNA sequences and RNA transcripts thereof) encoding a novel human polypeptide, "AFM" as well as complexes of AFM with ApoA1 and/or lipids, and polypeptide variants (including fragments and analogs) thereof which display  
35 one or more biological activities or properties specific to AFM.



Preferred DNA sequences of the invention include genomic and cDNA sequences as well as wholly or partially chemically and or enzymatically synthesized DNA sequences and biological replicas thereof. Also  
5 provided are autonomously replicating recombinant constructions such as plasmid and viral DNA vectors incorporating such sequences and especially vectors, wherein DNA encoding AFM or an AFM variant are operatively linked to an endogenous or exogenous  
10 expression control DNA sequence.

According to another aspect of the invention, host cells, especially unicellular host cells such as prokaryotic and eukaryotic cells, are stably transformed with DNA sequences of the invention in a manner allowing  
15 AFM and variants thereof to be expressed therein.

Host cells of the invention are useful in methods for the large scale production of AFM and AFM variants wherein the cells are grown in a suitable culture medium and the desired polypeptide products are  
20 isolated from the cells or from the growth medium.

Novel AFM and AFM variant products of the invention may be obtained as isolates from natural cell sources, but are preferably produced by recombinant procedures involving host cells of the invention. The  
25 products may be obtained in fully or partially glycosylated, partially or wholly deglycosylated, or non-glycosylated forms, depending on the host cell selected for recombinant production and/or post-isolation processing. The products may also be bound to  
30 other molecules, such as cellularly derived lipids and/or ApoA1.

Products of the invention include monomeric and multimeric polypeptides having the sequence of amino acid residues numbered -21 through 578 as set out in  
35 FIG. 1 herein. As explained in detail *infra*, this sequence includes a putative signal or leader sequence

which precedes the "mature" protein sequence and spans residue -21 (Met) through residue -1 (Thr) followed by the mature protein spanning residues 1 (Leu) to residue 578 (Asn). Based on amino acid composition, the  
5 calculated molecular weight of the mature protein lacking glycosylation or other post-translational modification is approximately 66,576 daltons.

AFM variants of the invention may comprise fragments including one or more of the regions specified  
10 herein and may also comprise polypeptide analogs wherein one or more of the specified amino acids is deleted or replaced: (1) without substantial loss, and preferably with enhancement, of one or more biological activities or immunological characteristics specific for AFM; or  
15 (2) with specific modulation of a particular ligand/receptor binding function. Analog polypeptides including additional amino acid residues (e.g., lysine) that facilitate multimer formation are also contemplated.

20 Further comprehended by the present invention are antibodies (e.g., monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, CDR-grafted antibodies and the like) or other binding proteins which are specific for AFM or AFM  
25 variants. Antibodies can be developed using isolated natural or recombinant AFM or AFM variants.

The antibodies are useful in complexes for immunization as well as for purifying polypeptides of the invention. The antibodies are also useful in  
30 modulating (i.e., blocking, inhibiting or stimulating) ligand/receptor binding reactions involving AFM.

Anti-idiotypic antibodies specific for anti-AFM antibodies and uses of such anti-idiotypic antibodies in treatment are also contemplated. Assays  
35 for the detection and quantification of AFM on cell surfaces and in fluids such as serum may involve a

single antibody or multiple antibodies in a "sandwich" assay format.

The uses of the DNA and amino acid sequences of the present invention are varied. For example, knowledge of the sequence of a cDNA for AFM makes possible the isolation by DNA/DNA hybridization of genomic DNA sequences encoding AFM and specifying AFM expression control regulatory sequences such as promoters, operators and the like. DNA/DNA hybridization procedures carried out with DNA sequences of the invention and under stringent conditions are likewise expected to allow the isolation of DNAs encoding allelic variants of AFM, other structurally related proteins sharing the biological and/or immunological specificity of AFM, and proteins homologous to AFM from non-human species (especially from other mammals). DNAs of the invention are useful in DNA/RNA hybridization assays to detect the capacity of cells to synthesize AFM. A variety of specific uses for AFM are disclosed herein below. These uses are primarily based on the known uses of the homologous albumin type polypeptides discussed above.

Also made available by the invention are anti-sense polynucleotides (e.g., DNA and RNA) relevant to regulating expression of AFM by those cells which ordinarily express it. Furthermore, knowledge of the DNA and amino acid sequences of AFM make possible the generation by recombinant means of hybrid fusion proteins characterized by the presence of AFM protein sequences and immunoglobulin heavy chain constant regions and/or hinge regions. See, Capon, et al., *Nature*, **337**: 525-531 (1989); Ashkenazi, et al., *P.N.A.S. (USA)*, **88**: 10535-10539 (1991); and *PCT WO 89/02922*, published April 6, 1989.

## BRIEF DESCRIPTION OF THE FIGURES

Numerous other aspects and advantages of the present invention will therefore be apparent upon  
5 consideration of the following detailed description thereof, reference being made to the drawings wherein:

FIG. 1 shows the nucleotide and deduced amino acid sequence of AFM. The putative signal sequence is indicated in lower case letters. Asterisks indicate  
10 putative sites for N-glycosylation. These are also represented as SEQ ID NO:1 AND SEQ ID NO:2.

FIG. 2A and 2B show a comparison of ALB family amino acid sequences. FIG. 2A shows the alignment of ALB family proteins. Sequences were aligned using the  
15 Clustal method in the MegAlign program (DNASTAR). Identical amino acid residues are boxed. Consensus indicates residues identical in all 4 sequences. Majority indicates 2 or 3 residues identical in all 4 sequences. FIG. 2B shows percent similarity (right of  
20 diagonal) and identity (left of diagonal) between ALB family members. Similarities were determined using the GCG GAP program. The sequences for the comparison proteins are also provided as follows: serum albumin, SEQ ID NO:3; alpha fetoprotein, SEQ ID NO:4; and vitamin  
25 D binding protein, SEQ ID NO:5.

FIG. 3 shows the conserved Cys pattern in ALB family proteins. The mature form of ALB family proteins are depicted with thin vertical bars representing single Cys residues and thick vertical lines representing -Cys-  
30 Cys- sequences.

FIG. 4 shows the putative disulfide bonding pattern for AFM. The organization of domains and double loops are drawn as originally proposed for ALB.

FIG. 5 shows the expression of rAFM in stably transfected CHO D<sup>-</sup> cells. Samples were applied to SDS-PAGE under reducing conditions followed by electrophoretic transfer to nitrocellulose. After blocking with skim milk, the membrane was probed with the AM339 antibody followed by incubation with donkey rabbit anti-Ig. Immunoreactive proteins were visualized by chemiluminescence. Lane 1, 80 ml conditioned medium (CM) from CHO D<sup>-</sup> cells transfected with AFM cDNA; lane 2, 80 ml CM from nontransfected CHO D<sup>-</sup> cells; lane 3, 100 ng AFM purified from human plasma. Size markers (in kDa) are indicated on the left.

FIG. 6 shows purification of rAFM. Samples purified from CHO D<sup>-</sup> cells transfected with AFM cDNA were adjusted with an equal volume of 2X sample buffer (125 mM Tris-HCl, pH 6.8, 4% SDS, 0.005% bromophenol blue, 10% glycerol) and analyzed by SDS-PAGE using 4-20% polyacrylamide gradient gels (Novex) under reducing conditions. The gel was stained with Coomassie Brilliant Blue. Lanes 1 and 6, 10 mg Mark-12 molecular weight markers (Novex, Inc.); lane 2, 50 mg total protein after addition of ammonium sulfate to concentrated CM (supernatant was dialyzed against PBS and subsequently loaded onto the gel); lane 3, 25 mg Phenyl Sepharose water eluate; lane 4, 10 mg Q Sepharose-purified rAFM; lane 5, 1 mg of Superdex 200-purified rAFM.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

### Definition of AFM

AFM is defined as a polypeptide having a qualitative biological activity or property in common with AFM of FIG. 1.

Included within the scope of AFM as that term is used herein is the AFM having the amino acid sequence of AFM as set forth in FIG. 1, SEQ ID NO:2; glycosylated, deglycosylated or unglycosylated derivatives of AFM; and lipidated or delipidated forms of AFM.

Also included within the scope of AFM are AFMs from any species, including without limitation: human, mouse, rat, pig, rabbit, monkey, dog, etc. Especially preferred is the human form of AFM.

#### Variants of AFM

Variants of AFM include homologous amino acid sequence variants of the sequence of FIG. 1, and homologous in-vitro-generated variants and derivatives of AFM, which are capable of exhibiting a biological activity or property in common with AFM of FIG. 1.

"Homologous" is used herein to refer to the residues in a candidate sequence that are identical with the residues in the sequence of AFM in FIG. 1 after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology.

A biological activity or property of AFM is defined as either 1) immunological cross-reactivity with at least one epitope of AFM, or 2) the possession of at least one regulatory or effector function qualitatively in common with AFM. Examples of those activities may be found in the section herein describing uses of AFM.

"Immunologically cross-reactive" as used herein means that the candidate polypeptide is capable of competitively inhibiting the qualitative biological activity of AFM or an AFM variant having this activity with polyclonal antisera raised against the known active analog. Such antisera are prepared in conventional fashion by injecting animals such as goats or rabbits, for example, subcutaneously with the known active analog

in complete Freund's adjuvant, followed by booster intraperitoneal or subcutaneous injection in incomplete Freund's adjuvant. An example of production of polyclonal antisera production is presented in the examples section below.

Amino acid sequence variants of AFM are prepared with various objectives in mind, including increasing the affinity of AFM for its binding partner, facilitating the stability, purification and preparation of AFM, and the like.

Amino acid sequence variants of AFM fall into one or more of three classes: insertional, substitutional, or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding AFM, by which DNA encoding the variant is obtained, and thereafter expressing the DNA in recombinant cell culture. However, variant AFM fragments having up to about 100-150 amino acid residues are prepared conveniently by in vitro synthesis.

The amino acid sequence variants of AFM are predetermined variants not found in nature or naturally occurring alleles. AFM variants typically exhibit the same qualitative biological activity as the naturally occurring AFM molecule. However, AFM variants and derivatives that are not capable of binding to their ligands are useful nonetheless (a) as a reagent in diagnostic assays for AFM or antibodies to AFM, (b) when insolubilized in accordance with known methods, as agents for purifying anti-AFM antibodies from antisera or hybridoma culture supernatants, and (c) as immunogens for raising antibodies to AFM or as immunoassay kit components (labeled, as a competitive reagent for the native AFM or unlabeled as a standard for AFM assay) so long as at least one AFM epitope remains active.

While the site for introducing an amino acid sequence variation is predetermined, the mutation per se need not be predetermined. For example, in order to optimize the performance of a mutation at a given site, random or saturation mutagenesis (where all 20 possible residues are inserted) is conducted at the target codon and the expressed AFM variant is screened for the optimal combination of desired activities. Such screening is within the ordinary skill in the art.

10 Amino acid insertions usually will generally be on the order of about from 1 to 10 amino acid residues; substitutions are typically introduced for single residues; and deletions will generally range about from 1 to 30 residues. Deletions or insertions  
15 preferably are made in adjacent pairs, i.e. a deletion of 2 residues or insertion of 2 residues. It will be amply apparent from the following discussion that substitutions, deletions, insertions or any combination thereof are introduced or combined to arrive at a final  
20 construct.

Insertional amino acid sequence variants of AFM are those in which one or more amino acid residues extraneous to AFM are introduced into a predetermined site in the target AFM and which displace the  
25 preexisting residues. Commonly, insertional variants are fusions of heterologous proteins or polypeptides to the amino or carboxyl terminus of AFM. Such variants are referred to as fusions of AFM and a polypeptide containing a sequence which is other than that which is  
30 normally found in AFM at the inserted position. Several groups of fusions are contemplated herein.

Immunologically active AFM derivatives and fusions comprise AFM and a polypeptide containing a non-AFM epitope, and are within the scope of this invention.  
35 The non-AFM epitope is any immunologically competent polypeptide, i.e., any polypeptide which is capable of



eliciting an immune response in the animal to which the fusion is to be administered or which is capable of being bound by an antibody raised against the non-AFM polypeptide. Typical non-AFM epitopes will be those  
5 which are borne by allergens, autoimmune epitopes, or other potent immunogens or antigens recognized by pre-existing antibodies in the fusion recipient, including bacterial polypeptides such as trpLE, beta-galactosidase, viral polypeptides such as herpes gD  
10 protein, and the like.

Immunogenic fusions are produced by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding an immunogenic polypeptide. It is preferable that the immunogenic  
15 fusion be one in which the immunogenic sequence is joined to or inserted into AFM or fragment thereof by a peptide bond(s). These products therefore consist of a linear polypeptide chain containing AFM epitope and at least one epitope foreign to AFM. It will be understood  
20 that it is within the scope of this invention to introduce the epitopes anywhere within the AFM molecule or fragment thereof.

Such fusions are conveniently made in recombinant host cells or by the use of bifunctional  
25 cross-linking agents. The use of a cross-linking agent to fuse AFM to the immunogenic polypeptide is not as desirable as a linear fusion because the cross-linked products are not as easily synthesized in structurally homogeneous form.

30 These immunogenic insertions are particularly useful when formulated into a pharmacologically acceptable carrier and administered to a subject in order to raise antibodies against AFM, which antibodies in turn are useful in diagnostics or in purification of  
35 AFM by immunoaffinity techniques known per se. In diagnostic applications, the antibodies will typically

be bound to or associated with a detectable group, examples of which are well known to those skilled in the art. Immunoaffinity techniques could be used, for example, to purify AFM.

5           Other fusions, which may or may not also be immunologically active, include fusions of the mature AFM sequence with a signal sequence heterologous to AFM, and fusions of AFM to polypeptides having enhanced plasma half life (ordinarily > about 20 hours) such as  
10 immunoglobulin chains or fragments thereof.

          Signal sequence fusions are employed in order to more expeditiously direct the secretion of AFM. The heterologous signal replaces the native AFM signal, and when the resulting fusion is recognized, i.e. processed  
15 and cleaved by the host cell, AFM is secreted. Signals are selected based on the intended host cell, and may include bacterial yeast, mammalian and viral sequences. The native AFM signal or the herpes gD glycoprotein signal is suitable for use in mammalian expression  
20 systems.

          Substantial variants are those in which at least one residue in the FIG. 2 sequence has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with  
25 the following Table 1 when it is desired to finely modulate the characteristics of AFM.

TABLE 1

Original Residue	Exemplary Substitutions
Ala	Ser
Arg	Lys
Asn	Gln; His
Asp	Glu
Cys	Ser; Ala
Gln	Asn
Glu	Asp
Gly	Pro
His	Asn; Gln
Ile	Leu; Val
Leu	Ile; Val
Lys	Arg; Gln; Glu
Met	Leu; Ile
Phe	Met; Leu; Tyr
Ser	Thr
Thr	Ser
Trp	Tyr
Tyr	Trp; Phe
Val	Ile; Leu

Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in

5 Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical

10 molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in AFM properties will

be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g. glycine.

Some deletions, insertions, and substitutions will not produce radical changes in the characteristics of AFM molecule. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, for example when modifying the AFM extracellular domain or an immune epitope, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays.

Another class of AFM variants are deletional variants. Deletions are characterized by the removal of one or more amino acid residues from AFM sequence. Deletions from the AFM C-terminal or the N-terminal, which preserve the biological activity or immune cross-reactivity of AFM are suitable.

Deletions of cysteine or other labile residues also may be desirable, for example in increasing the oxidative stability of AFM. Deletion or substitutions of potential proteolysis sites, e.g. Arg-Arg, is accomplished by deleting one of the basic residues or substituting one by glutamyl or histidyl residues.

Preferably, the variants represent conservative substitutions. It will be understood that some variants may exhibit reduced or absent biological activity. These variants nonetheless are useful as

standards in immunoassays for AFM so long as they retain at least one immune epitope of AFM.

Glycosylation variants are included within the scope of AFM. They include variants completely lacking  
5 in glycosylation (nonglycosylated) and variants having at least one less glycosylated site than the native form (deglycosylated) as well as variants in which the glycosylation has been changed. Additionally, unglycosylated AFM which has the amino acid sequence of  
10 the native AFM is produced in recombinant prokaryotic cell culture because prokaryotes are incapable of introducing glycosylation into polypeptides.

Glycosylation variants are produced by selecting appropriate host cells or by in vitro methods.  
15 Yeast, for example, introduce glycosylation which varies significantly from that of mammalian systems. Similarly, mammalian cells having a different species (e.g. hamster, murine, insect, porcine, bovine or ovine) or tissue origin (e.g. lung, lymphoid, mesenchymal or  
20 epidermal) than the source of AFM are routinely screened for the ability to introduce variant glycosylation as characterized for example by elevated levels of mannose or variant ratios of mannose, fucose, sialic acid, and other sugars typically found in mammalian glycoproteins.  
25 In vitro processing of AFM typically is accomplished by enzymatic hydrolysis, e.g. neuraminidase digestion.

AFM isolated from natural sources or produced recombinantly will generally contain bound lipids. The nature of the bound lipids is expected to depend on the  
30 source of the AFM. Delipidated versions of AFM may be prepared by standard delipidation methods known in the art, especially the art relating to ALB where delipidation is a common procedure. One preferred method involves extracting an aqueous solution of AFM  
35 with a solvent capable of dissolving lipids, such as 1-

butanol or diisopropyl ether. Example 8 presents a specific exemplary method for delipidation.

Certain post-translational derivatizations are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains (T. E. Creighton, *Proteins: Structure and Molecular Properties*, (W. H. Freeman & Co.), San Francisco: 79-86 (1983), acetylation of the N-terminal amine and, in some instances, amidation of the C-terminal carboxyl.

20

#### Oligonucleotides relating to AFM

DNA encoding AFM is synthesized by in vitro methods or is obtained readily from human liver cDNA libraries. The means for synthetic creation of the DNA encoding AFM, either by hand or with an automated apparatus, are generally known to one of ordinary skill in the art, particularly in light of the teachings contained herein. As examples of the state of the art relating to polynucleotide synthesis, one is directed to Maniatis et al., *Molecular Cloning-A Laboratory Manual*, Cold Spring Harbor Laboratory (1984), and Horvath et al., *An Automated DNA Synthesizer Employing Deoxynucleoside 3'-Phosphoramidites*, *Methods in Enzymology* 154: 313-326 (1987).

Alternatively, to obtain DNA encoding AFM from sources other than murine or human, since the entire DNA

sequence for the preferred embodiment of AFM is given, one needs only to conduct hybridization screening with labelled DNA encoding AFM or fragments thereof (usually, greater than about 20, and ordinarily about 50bp) in order to detect clones which contain homologous sequences in the cDNA libraries derived from the liver of the particular animal, followed by analyzing the clones by restriction enzyme analysis and nucleic acid sequencing to identify full-length clones. If full length clones are not present in the library, then appropriate fragments are recovered from the various clones and ligated at restriction sites common to the fragments to assemble a full-length clone. DNA encoding AFM from other animal species is obtained by probing libraries from such species with the human sequences, or by synthesizing the genes in vitro.

Included within the scope hereof are nucleic acid sequences that hybridize under stringent conditions to a fragment of the DNA sequence in FIG. 1, which fragment is greater than about 10 bp, preferably 20-50 bp, and even greater than 100 bp. Also included within the scope hereof are nucleic acid sequences that hybridize under stringent conditions to a fragment of AFM. "Stringent" is used to refer to conditions that are commonly understood in the art as stringent. An exemplary set of conditions include a temperature of 60 - 70°C, (preferably about 65°C) and a salt concentration of 0.70M to 0.80M (preferably about 0.75M).

Included also within the scope hereof are nucleic acid probes which are capable of hybridizing under stringent conditions to the cDNA of AFM or to the genomic gene for AFM (including introns and 5' or 3' flanking regions extending to the adjacent genes or about 5,000 bp, whichever is greater).

### Recombinant Expression of AFM

In general, prokaryotes are used for cloning of DNA sequences in constructing the vectors useful in the invention. For example, *E. coli* K12 strain 294  
5 (ATCC No. 31446) is particularly useful. Other microbial strains which may be used include *E. coli* B and *E. coli* X1776 (ATCC No. 31537). Alternatively, in vitro methods of cloning, e.g. polymerase chain reaction, are suitable.

10 AFMs of this invention are expressed directly in recombinant cell culture as an N-terminal methionyl analog, or as a fusion with a polypeptide heterologous to AFM, preferably a signal sequence or other polypeptide having a specific cleavage site at the N-  
15 terminus of AFM. For host prokaryotes that do not process AFM signal, the signal is substituted by a prokaryotic signal selected for example from the group of the alkaline phosphatase, penicillinase, or heat stable enterotoxin II leaders. For yeast secretion the  
20 human AFM signal may be substituted by the yeast invertase, alpha factor or acid phosphatase leaders. In mammalian cell expression the native signal is satisfactory for mammalian AFM, although other mammalian secretory protein signals are suitable, as are viral  
25 secretory leaders, for example the herpes simplex gD signal.

AFM may be expressed in any host cell, but preferably are synthesized in mammalian hosts. However, host cells from prokaryotes, fungi, yeast, insects and  
30 the like are also used for expression. Exemplary prokaryotes are the strains suitable for cloning as well as *E. coli* W3110 (F-, l-, prototrophic, ATCC No. 7325), other enterobacteriaceae such as *Serratia marcescens*, bacilli and various pseudomonads. Preferably the host  
35 cell should secrete minimal amounts of proteolytic enzymes.



Expression hosts typically are transformed with DNA encoding AFM which has been ligated into an expression vector. Such vectors ordinarily carry a replication site (although this is not necessary where  
5 chromosomal integration will occur). Expression vectors also include marker sequences which are capable of providing phenotypic selection in transformed cells, as will be discussed further below. Expression vectors also optimally will contain sequences which are useful  
10 for the control of transcription and translation, e.g., promoters and Shine-Dalgarno sequences (for prokaryotes) or promoters and enhancers (for mammalian cells). The promoters may be, but need not be, inducible.

Promoters suitable for use with prokaryotic  
15 hosts illustratively include the  $\beta$ -lactamase and lactose promoter systems (Chang et al., *Nature*, **275**, 615 (1978); and Goeddel et al., *Nature* **281**, 544 (1979)), alkaline phosphatase, the tryptophan (trp) promoter system (Goeddel, *Nucleic Acids Res.* **8**: 4057 (1980) and EPO  
20 Appln. Publ. No. 36,776) and hybrid promoters such as the tac promoter (H. de Boer et al., *Proc. Natl. Acad. Sci. USA* **80**, 21-25 (1983)). However, other functional bacterial promoters are suitable. Their nucleotide sequences are generally known, thereby enabling a  
25 skilled worker operably to ligate them to DNA encoding AFM (Siebenlist et al., *Cell* **20**, 269 (1980)) using linkers or adapters to supply any required restriction sites. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked  
30 to the DNA encoding AFM.

In addition to prokaryotes, eukaryotic microbes such as yeast or filamentous fungi are satisfactory. *Saccharomyces cerevisiae* is the most commonly used eukaryotic microorganism, although a  
35 number of other strains are commonly available.

Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., *J. Biol. Chem.* **255**, 2073 (1980)) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Reg.* **7**: 149 (1968); and Holland, *Biochemistry* **17**, 4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde 3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in R. Hitzeman et al., *European Patent Publication No. 73,657A*.

Expression control sequences are known for eukaryotes. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide. At the 3' end of most eukaryotic genes is an AATAAA sequence which may be the signal for addition of the poly A tail to the 3' end of the coding sequence. All of these sequences are inserted into mammalian expression vectors.

Suitable promoters for controlling transcription from vectors in mammalian host cells are readily obtained from various sources, for example, the

genomes of viruses such as polyoma virus. SV40, adenovirus, MMV (steroid inducible), retroviruses (e.g. the LTR of HIV), hepatitis-B virus and most preferably cytomegalovirus, or from heterologous mammalian

5 promoters, e.g. the beta actin promoter. The early and late promoters of SV40 are conveniently obtained as an SV40 restriction fragment which also contains the SV40 viral origin of replication. Fiers et al., *Nature* 273, 113 (1978). The immediate early promoter of the human

10 cytomegalovirus is conveniently obtained as a HindIII E restriction fragment. Greenaway, P. J. et al., *Gene* 18, 355-360 (1982).

Transcription of a DNA encoding AFM by higher eukaryotes is increased by inserting an enhancer

15 sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10-300 bp, that act on a promoter to increase its transcription. Enhancers are relatively orientation and position independent having been found 5' (Laimins, L. et al., *PNAS* 78, 993

20 (1981)) and 3' (Lusky, M. L., et al., *Mol. Cell Bio.* 3, 1108 (1983) to the transcription unit, within an intron (Banerji, J. L. et al., *Cell* 33: 729 (1983)) as well as within the coding sequence itself (Osborne, T. F., et al., *Mol. Cell Bio.* 4, 1293 (1984)). Many enhancer

25 sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp

30 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human or

35 nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of

transcription which may affect mRNA expression. These regions are transcribed as polyadenylated segments in the untranslated portion of the mRNA encoding AFM. The 3' untranslated regions also include transcription termination sites.

Expression vectors may contain a selection gene, also termed a selectable marker. Examples of suitable selectable markers for mammalian cells are dihydrofolate reductase (DHFR), thymidine kinase (TK) or neomycin.

Suitable eukaryotic host cells for expressing AFM include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham, F. L. et al., J. Gen Virol. 36: 59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary-cells-DHFR (CHO, Urlaub and Chasin, PNAS (USA) 77, 4216, (1980)); mouse sertoli cells (TM4, Mather, J. P., Biol. Reprod. 23, 243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); and, TRI cells (Mather, J. P. et al., Annals N.Y. Acad. Sci. 383, 44-68 (1982)).

Construction of suitable vectors containing the desired coding and control sequences employ standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required.

Host cells are transformed with the expression vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants or amplifying AFM

gene. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

5           The host cells referred to in this disclosure encompass cells in in vitro culture as well as cells which are within a host animal.

          "Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an  
10 extrachromosomal element or by chromosomal integration. The method used herein for transformation of the host cells may be, for example, the method of Graham, F. and van der Eb, A., *Virology* 52, 456-457 (1973). However, other methods for introducing DNA into cells such as by  
15 nuclear injection or by protoplast fusion may also be used. If prokaryotic cells or cells which contain substantial cell wall constructions are used, the preferred method of transfection is calcium treatment using calcium chloride as described by Cohen, F. N.  
20 et al., *Proc. Natl. Acad. Sci. (USA)* 69, 2110 (1972).

          "Transfection" refers to the introduction of DNA into a host cell whether or not any coding sequences are ultimately expressed. Numerous methods of  
transfection are known to the ordinarily skilled  
25 artisan, for example,  $\text{CaPO}_4$  and electroporation. Transformation of the host cell is indicative of successful transfection.

#### Recovery and Purification of AFM

30           AFM is recovered and purified from recombinant cell cultures by known methods, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, immunoaffinity chromatography,  
35 hydroxyapatite chromatography, lectin chromatography, hydrophobic interaction chromatography, and gel

filtration chromatography. Other known purification methods within the scope of this invention utilize, for example, immobilized carbohydrates, epidermal growth factor, or complement domains. Moreover, reverse-phase HPLC and chromatography using anti-AFM antibodies are useful for the purification of AFM. AFM may preferably be purified in the presence of a protease inhibitor such as PMSF. A specific preferred method of purifying AFM is found in Example 1.

10

#### Uses of AFM

As mentioned above, in view of the homology between AFM and each of ALB, AFP and VDB, it is expected that AFM and variants thereof will have identical or similar biological activities and utilities. In general, members of the albumin family show a high propensity to bind and transport a wide variety of substances through the body including fatty acids, hormones, enzymes, dyes, trace metals and drugs. It is expected that a preparation of AFM which has its endogenous lipids removed could be used to reduce high concentrations of free fatty acids (hyperlipidemia) found in disease states such as acute pancreatitis, ARDS, sepsis and atherosclerosis.

Similar to proposed functions for ALB, AFM can also potentially be used as an antioxidant. AFM is expected to act as an antioxidant by sequestering metal ions and preventing those ions from accelerating free-radical reactions (ie. decomposition of lipid peroxides to peroxy and alkoxy radicals; formation of hydroxyl radical from hydrogen peroxide) or AFM may act to directly inactivate hydrogen peroxide, hydroxyl radicals and hypochlorous acid. The detoxification of oxygen metabolites by AFM can limit the detrimental effects that unbound oxygen metabolites have in inactivating beneficial anti-proteases ( $\alpha_1$ -antiproteinase and  $\alpha_1$ -

antitrypsin) and in damaging DNA, proteins and lipids. Thus, AFM may be used to ameliorate ischaemia-reperfusion injury, rheumatoid arthritis, ARDS, cardiopulmonary bypass, sepsis and any other diseases or  
5 tissue damage caused by an excess production of oxygen metabolites from leukocytes and/or arachidonic acid metabolism.

Another use of AFM is based on the finding that albumin family proteins bind and detoxify a wide  
10 variety of toxic products including man-made drugs. Thus, AFM can ameliorate the effects of toxic plasma substances released as a result of inflammation and can be conjugated to toxic pharmaceuticals so as to minimize detrimental activities of the drug.

15 AFM may also have anticoagulant properties as has been reported for ALB. AFM may inhibit platelet aggregation and thus could be a useful additive to resuscitation fluids in disease states characterized by increased platelet aggregation such as sepsis,  
20 hemmorrhagic shock and burn injury.

Serum AFM levels may increase or decrease due to a particular pathological condition as has been observed with AFP. Thus, calculation of serum AFM levels using antibodies described in Example 11 will be useful  
25 in diagnosing specific human diseases.

#### Administration of AFM

For administration in vivo, AFM is placed into sterile, isotonic formulations, and administered by  
30 standard means well known in the field. The formulation of AFM is preferably liquid, and is ordinarily a physiologic salt solution containing 0.5-10 mM calcium, non-phosphate buffer at pH 6.8-7.6, or may be lyophilized powder.

35 It is envisioned that intravenous delivery, or delivery through catheter or other surgical tubing will

be the primary route for therapeutic administration. Alternative routes include tablets and the like, commercially available nebulizers for liquid formulations, and inhalation of lyophilized or aerosolized receptors. Liquid formulations may be utilized after reconstitution from powder formulations.

AFM may also be administered via microspheres, liposomes, other microparticulate delivery systems or sustained release formulations placed in certain tissues including blood.

The dose of AFM administered will be dependent upon the properties of AFM employed, e.g. its activity and biological half-life, the concentration of AFM in the formulation, the administration route for AFM, the site and rate of dosage, the clinical tolerance of the patient involved, the pathological condition afflicting the patient and the like, as is well within the skill of the physician.

AFM may also be administered along with other pharmacologic agents such as antibiotics, anti-inflammatory agents, and anti-tumor agents. It may also be useful to administer AFM along with other therapeutic proteins such as gamma-interferon and other immunomodulators.

25

#### Ligands of AFM

It is also possible that the AFM of the present invention will have an additional natural ligand or ligands (other than the lipids discussed herein) and that such ligands will be capable of modulating the biological activities of AFM in vivo. One of ordinary skill may be able to utilize the teachings of the present invention to screen for ligands of the AFMs disclosed herein and then isolate and purify them, e.g., by immunochromatography using AFM or a variant of AFM bound to a solid support.



Antibodies to AFM

Also included within the scope of the invention are antibodies (e.g., monoclonal and  
5 polyclonal antibodies, single chain antibodies, chimeric antibodies, CDR-grafted antibodies and the like) or other binding proteins which are specific for AFM or its variants. Antibodies can be developed using natural or recombinant AFM or AFM variants or cells expressing such  
10 molecules on their surfaces. Active fragments of such antibodies are also contemplated.

The antibodies are useful for purifying the polypeptides described herein, or identifying and purifying cells producing the polypeptides on their  
15 surfaces. The antibodies could also be used to modulate (e.g., block, inhibit or stimulate) ligand binding to AFM. Anti-idiotypic antibodies are also contemplated. Assays for detection and quantitation of AFM on cell surfaces and in fluids (such as serum, blood,  
20 cerebrospinal fluid, urine, semen, milk, and tears) may involve a single antibody or a "sandwich" assay format.

\* \* \*

25 It is understood that the application of the teachings of the present invention to a specific problem or situation will be within the capabilities of one having ordinary skill in the art in light of the teachings contained herein. Examples of the products of  
30 the present invention and representative processes for their isolation, use, and manufacture appear below. All literature and patent citations herein are expressly incorporated by reference.

## EXAMPLES

EXAMPLE 1: Lipopolysaccharide (LPS) Neutralization Assay

## 5 A. Formation of ELPS.

Sheep erythrocytes are first coated with *S. minnesota* Re595 LPS (List Biological) as previously described (*J. Ex. Med.* 170:1231). The concentration of LPS used to coat erythrocytes is 10 mg/8 x 10<sup>7</sup> erythrocytes. After coating with LPS, erythrocytes are resuspended in 5 mM veronal buffer, pH 7.5, containing 150 mM NaCl, 0.1% gelatin and 1 mM EDTA (EDTAGVB<sup>2-</sup>) to a concentration of 1x10<sup>8</sup> cells/ml.

## 15 B. Opsonization of ELPS.

ELPS are opsonized using normal human plasma (NHP) as a source of opsonizing proteins (septin) (*J. Ex. Med.* 176:719) or LBP (*J. Ex. Med.* 170:1231). NHP is diluted 1:100 in phosphate buffered saline without divalent cations containing 1 mM EDTA (PDEDTA) and then mixed 1:1 with ELPS (1 x 10<sup>8</sup> cells/ml). The mixture is then incubated 10 minutes at 37°C, spun down for 1 minute at 800x g in a swinging bucket centrifuge. Cells are washed 2x with EDTAGVE<sup>2-</sup> and resuspended to a final concentration of 1 x 10<sup>8</sup> cells/ml in EDTAGVB<sup>2-</sup>. Erythrocytes treated as above are denoted E-septins.

## C. Assay of Samples for LPS Neutralizing Activity.

The following step measures the ability of a given sample to reduce the amount of binding of LPS coated erythrocytes to macrophages in a CD-14 dependent manner.

Dilutions of samples to be tested are made in PDEDTA. E-septins are incubated 1:1 with sample dilutions at 37°C for 40 minutes. One sample is included in which E-septin is incubated with PDEDTA alone as a control. Following incubation, E-septins are

spun down in a swinging bucket centrifuge at 800x g for 1 minute and resuspended in EDTAGVB<sup>2-</sup> to a concentration of  $1 \times 10^8$  cells/ml. E-septins treated in this manner are denoted E-septin<sup>1</sup>. Five ml of the resuspended E-septins<sup>1</sup> are added to a macrophage monolayer (see below) in a 60 well terasaki plate (NUNC, Inc.). The plate is incubated 20 minutes at room temperature followed by inversion and incubation another 12 minutes at room temperature to separate unbound E-septins<sup>1</sup> from the macrophages. The plate is then dipped in a beaker of cold PBS with no divalent cations containing 0.001% azide to wash off unbound erythrocytes. Binding of E-septins<sup>1</sup> is evaluated by phase contrast microscopy. Binding is expressed as the attachment index which denotes the number of erythrocytes bound per 100 macrophages. LPS neutralizing activity of the samples is evaluated by comparing the binding of E-septins<sup>1</sup> incubated with dilutions of the sample in PDEDTA to E-septin<sup>1</sup> incubated with PDEDTA alone. Activity is measured in units/ml. This represents the dilution at which the sample reduces E-septin<sup>1</sup> binding by 50% in the standard assay.

#### D. Macrophage Monolayer.

Monocyte derived macrophages are obtained by culturing human monocytes in teflon beakers as described (*J. Ex. Med.* 156:1149). On the day of the assay, macrophages are removed from the beaker, washed with PBS containing divalent cations (PBS<sup>2+</sup>) and resuspended in PBS<sup>2+</sup> containing 0.5 mg/ml human serum albumin (Armour Pharm.), 0.3 U/ml aprotinin (Sigma) and 3 mM glucose (HAP buffer) to  $1 \times 10^6$  cells/ml. Thirty minutes prior to addition of E-septins<sup>1</sup> to the terasaki plate, macrophage monolayers are formed by adding 5 ml of cells and 5 ml of HAP buffer to each well. Immediately prior

to addition of E-septins<sup>1</sup>, the plate is flooded with PDEDTA and the wells are lightly aspirated.

EXAMPLE 2: Purification of AFM

5 All steps, except where noted were done at controlled room temperature. All buffers were made with pyrogen free water and sterile filtered prior to use. One unit of fresh frozen plasma (FFP) was thawed by immersion of the packet in a room-temperature water  
10 bath. The thawed plasma was first transferred to a graduated cylinder, the volume noted, and then poured into a beaker. The FFP was then stirred magnetically while the beaker was immersed in a ice-water bath. To the FFP, sufficient 3.9M ammonium sulfate (4°C) was  
15 slowly added to achieve a final ammonium sulfate concentration of 1.6M. This mixture was allowed to stir at 4°C for an additional 4 hours. The mixture was then centrifuged at 10,000 x g for 60 minutes. The supernatant was recovered and allowed to come to room  
20 temperature by immersion in a room temperature water bath.

A portion (250 ml) of the ammonium sulfate supernatant was loaded onto a 2.6 x 10cm Phenyl Sepharose HP column previously equilibrated with 50mM  
25 sodium phosphate, 1.6M ammonium sulfate; pH 7.5. Loading, wash and elution steps were done at a linear flow rate of 40cm/hr. Following sample loading the column was washed in turn with 50mM sodium phosphate, 1.6M ammonium sulfate; pH 7.5 and 50mM sodium phosphate;  
30 pH 7.5. Elution of activity (see Example 1) was accomplished by the introduction of water. The pH of the water eluate was adjusted to 8.2 with 0.1M Tris base and loaded, at a flow rate of 150cm/hr, onto a 2.6 x 10cm Q Sepharose HP column previously equilibrated with  
35 20mM Tris-HCl; pH 8.2. Following sample loading the column was thoroughly washed with first 20mM Tris-HCl;

pH 8.2 and then 0.1M sodium chloride, 20mM Tris-HCl; pH 8.2. The column was then resolved with a 0.33mM/ml linear gradient of sodium chloride ranging from 0.1 to 0.3M, followed by a 7mM/ml linear gradient of sodium chloride ranging from 0.3M to 1.0M in 20mM Tris-HCl; pH 8.2. Active fractions were pooled (30 ml).

One third of the pooled activity from above was loaded onto a Superdex 200 prep column (2.6 x 60cm) previously equilibrated with phosphate buffered saline; pH 7.4. The column was loaded and resolved at a flow rate of 1ml/min. The fractions were analyzed for activity and by reducing SDS-PAGE. A pool (8ml) of the appropriate fractions showed two major protein bands by reducing SDS PAGE at a molecular weight of 87 kd and 28 kd. Two milliliters of the above pool was submitted for sequence analysis. The 28 kd protein had the same N-terminus as ApoA1 and the 87 kd protein had a novel N-terminus (see below). The remaining 6ml of the Superdex 200 prep pool was loaded onto a Superdex 200 prep column (2.6 x 60cm) previously equilibrated with 1% sodium deoxycholate, 0.15M sodium chloride, 50mM Tris-HCl; pH 8.5. The column was loaded and resolved at a flow rate of 1ml/min. The fractions were analyzed by SDS-PAGE. Those fractions corresponding to either the 87 kd or the 28 kd proteins were separately pooled (16ml each).

Four milliliters of each pool was diafiltered against 50mM Tris-HCl; pH 8.5 to remove excess sodium deoxycholate and concentrated to approximately 0.1mg/ml.

EXAMPLE 3: Amino Acid Sequence Analysis of N-terminus  
and Proteolytic Fragments of Purified AFM

To obtain the N-terminal sequence, 10 ug of purified AFM was loaded on the reverse phase portion of a precycled biphasic sequencing column. The column was loaded on a Hewlett Packard G1004A protein sequencer with on-line phenylthiohydantoin amino acid analysis performed with a Hewlett Packard 1090 high performance liquid chromatography (HPLC). The N-terminal amino acid sequence obtained from the protein was:

LPTQPRDIENFXSTQKFIEDNIEYITIIAFAQYVQ (SEQ ID NO: 6),

where "X" represents an unassignable amino acid.

To obtain sequences from AFM tryptic peptide fragments, 100 mg of AFM was dissolved in 8M urea with 0.4 M ammonium bicarbonate, then reduced with DTT and carboxymethylated with iodoacetic acid. The sample was subsequently diluted with water to adjust the urea concentration to 2M, then digested with sequencing grade trypsin (Boehringer Mannheim) at 37°C for 18 h with an enzyme to substrate ratio of 1:50. The digested protein was injected on a Hewlett Packard 1090 HPLC equipped with a 4.6 x 250 mm C18 reverse phase column (Vydac). Elution was performed using a linear gradient of acetonitrile with 0.1 % trifluoroacetic acid at a flow rate of 0.75 ml/min. Elution was monitored at 214 nm and fractions were collected. Selected fractions were run on the Hewlett Packard G1004A protein sequencer as described above. The amino acid sequences obtained from select tryptic fragments are summarized in Table 1.

Table 1: Trypsin fragments

	FX18	YHYLIR	(SEQ ID NO:7)
5	FX20	FTFEYSR	(SEQ ID NO:8)
	FX27	FTDSENVQCQERDADP	(SEQ ID NO:9)
	FX29	IVQIYKDILLR	(SEQ ID NO:10)
10	FX32	IAPQLSTEELVSLGE	(SEQ ID NO:11)
	FX37	RHPDLSIPELLR	(SEQ ID NO:12)
15	FX45	ESLLNHFLYEVAR	(SEQ ID NO:13)
	FX53	RNPFVFAPTLLTVAVHFEEVAKSCC	(SEQ ID NO:14)

20 EXAMPLE 4: Isolation of AFM cDNA

The polymerase chain reaction (PCR) was used to amplify a portion of the AFM gene from which an exact probe could be derived. A PCR (PCR 1) was first performed with fully degenerate primers specifying the sense strand for the N-terminal amino acid sequence QKFIEDN (SEQ ID NO: 15) [5' ACG CTG AAT TCG CCA (GA)AA (GA)TT (CT)AT (ATC)GA (GA)GA (CT)AA] (SEQ ID NO: 16) and the antisense strand for a portion of the FX 29 tryptic peptide sequence IVQIYKD (SEQ ID NO: 17) [5' ACG CTA AGC TTG C(GA)T C(CT)T T(GA)T A(GAT)A T(CT)T G(AGCT)A C(GAT)A T] (SEQ ID NO: 18). One ng of Quick Clone Human liver cDNA (Clontech, cat. no. 7113-1) was used as the template in a 100 ml PCR performed in a standard buffer (Perkin-Elmer Cetus) with 1 mM of each degenerate primer. The cycling parameters used in the PCR were as follows: 95°C , 8 min (1 cycle); 94°C, 1 min, 34°C, 10

min, 72°C, 2 min (3 cycles); 94°C, 1 min, 50°C, 1 min, 72°C, 2 min (45 cycles); 72°C, 5 min (1 cycle).

Agarose gel analysis of PCR 1 did not reveal the amplification of any specific products, thus we  
5 utilized an aliquot of PCR 1 as a template for a second PCR (PCR 2) using a nested primer pair. For this experiment, we used fully degenerate primers specifying the sense strand for the N-terminal amino acid sequence DNEYIT (SEQ ID NO: 19) [5' ACG CTG AAT TCG CGA (CT)AA  
10 (CT)AT (ATC)GA (GA)TA (CT)AT (ATC)AC] (SEQ ID NO: 20) and the antisense strand for the FX 20 tryptic peptide sequence FTFEYS (SEQ ID NO: 21) [5' ACG CTA AGC TTG C(GATC)G A(GA)T A(CT)T C(GA)A A(ACGT)G T(GA)A A] (SEQ ID NO: 22). PCR 2 was then performed using the same  
15 reaction mix and cycling parameters as PCR 1, except for the substitution of 1 ml of PCR 1 in place of the human liver cDNA. Analysis of PCR 2 by agarose gel electrophoresis revealed the amplification of a 1 kb product.

20 To prepare PCR 2 for DNA sequencing, the inventors utilized the *EcoR* I and *Hind* III sites that were incorporated into the degenerate primers. These sites were used to clone the fragment into mp19 (Boehringer Mannheim). The 1 kb PCR 2 product cloned in  
25 mp19 (mp19 AFM) was then sequenced in its entirety from both strands. Nucleotide sequence analysis of the fragment confirmed that we had amplified a segment of AFM cDNA as the sequence was found to have an open reading frame which encoded 3 tryptic peptides that were  
30 already sequenced (Table 1, Fxs 27, 45, 53). The nucleotide sequence of the 1 kb fragment was compared to all sequences in the Genbank database and found to be unique. We also observed that the partial AFM cDNA had significant homology with cDNAs reported for ALB family  
35 proteins (ALB, AFP and VDB). Therefore, in order to to isolate a full-length cDNA encoding AFM and to minimize



the probability of hybridizing other ALB family genes, we screened a human liver cDNA library with an exact 18mer oligonucleotide (5' TAT GTG CTA TGG AGG GGC) (SEQ ID NO: 23) derived from a segment of AFM sequence that was not highly homologous to ALB, AFP and VDB. The oligonucleotide was end-labeled with  $^{32}\text{P}_i$  and used to screen a human liver cDNA library (Clontech, cat. no. HL1115a). Positive clones were purified and rescreened with the same oligonucleotide probe. A single positive clone (17AFM) with a 2.3 kb insert was chosen for further study and phage DNA was prepared. The 2.3 kb insert was then sequenced in its entirety from both strands and verified to encode AFM.

This approach enabled the inventors to isolate a lambda phage (17AFM) containing full-length AFM cDNA.

#### EXAMPLE 5: Characterization of AFM

The insert in 17AFM is 2287 bp (FIG. 1) consisting of a 317 bp 5' untranslated region, a 1797 bp sequence encoding a protein of 599 amino acids and a 173 bp 3' untranslated region. The predicted amino acid sequence of AFM was found to contain all the tryptic peptides (See Table 1) that had been previously sequenced from the purified protein. Translation of the AFM cDNA sequence reveals that a 21 amino acid hydrophobic leader peptide precedes the experimentally determined N-terminus of mature AFM. The mature AFM is predicted to have 578 amino acids with a calculated Mr of 66576 and pI of 5.65. The difference between the calculated Mr of AFM and its apparent molecular weight on SDS-PAGE is likely due to glycosylation. AFM has 4 potential sites for N-glycosylation (FIG. 1) and N-glycanase treatment reduced the apparent Mr of AFM to 65000 when analyzed by reducing SDS-PAGE (data not shown).

A comparison of the deduced AFM amino acid sequence and other ALB family sequences is shown in FIG. 2. It is evident there is strong similarity between AFM and other ALB family members throughout the entire protein. When conserved amino acids are taken into account, the of AFM to AFP, ALB and VDB is 60.4%, 54.8% and 41.2% respectively. The distribution of Cys residues is conserved among ALB family members. The positions of AFM Cys residues are clearly consistent with this arrangement (FIG. 3). The Cys residues in ALB family proteins have also been proposed to form a pattern of disulfide bridges enabling these proteins to be depicted as a series of 9 double loops defining 3 structural domains. FIG. 4 shows that the 34 Cys residues in AFM can be arranged into a pattern of 17 disulfide-linked pairs that parallels the domain organization observed in other ALB family proteins.

#### EXAMPLE 6: Chromosomal Mapping

PCR was performed on a panel of somatic cell hybrids (Bios Laboratories, cat. no. CP2-02) to identify the chromosomal location of the gene. For PCR, we utilized primers (5' CAA CCC TGC TGT GGA CCA C; 5' GCA CAT ATG TTT TAT CAG CTT T) (SEQ ID NO: 24 and SEQ ID NO: 25) that would be expected to amplify an 88 bp fragment between nt 1790 and 1878 in the AFM cDNA. Each PCR on somatic cell hybrid DNA was performed in a standard 100 ml reaction mixture (Perkin-Elmer Cetus) containing 250 ng DNA and a final concentration of 0.1 mM of each primer. The cycling parameters were as follows: 95°C, 5 min (1 cycle); 94°C, 1 min, 56°C, 1 min, 72°C, 1 min (25 cycles); 72°C, 5 min (1 cycle).

Utilizing PCR on a panel of commercially available somatic cell hybrids (data not shown), we detected an amplified product in only 2 hybrids. Both of these hybrids had DNA in common from human chromosomes

4, 5 and 8. Since PCRs performed on hybrids containing DNA from chromosomes 5 and 8 did not yield amplified products, we conclude that AFM resides on human chromosome 4 along with other ALB family genes.

5

EXAMPLE 7: Stable Expression of rAFM

Two separate PCR's were performed on AFM cDNA to generate 2 overlapping fragments that span the entire AFM cDNA sequence. The oligonucleotide pair (5' TCA CCT CTA GAC CAC CAT GAA ACT ACT AAA ACT TAC AG + 5' AAT TTC TCA GGA GAT CTT TGT ATA) (SEQ ID NO: 26 and SEQ ID NO: 27) used in the first PCR introduced an *Xba* I site and a perfect Kozak sequence preceeding the AFM initiator codon. The amplified product was subsequently cloned into pGEMT (Promega) to create pDJ11. The oligonucleotide pair (5' AAA TAT ACA AAG ATC TCC TGA GAA + 5' TCC CGG TCG ACT CAG TTG CCA ATT TTT GGA C) (SEQ ID NO: 28 and SEQ ID NO: 29) used in the second PCR introduced a *Sal* I site following the natural stop codon in AFM and the product was cloned into pGEMT to create pDJ13. The 3' end of AFM was then joined to the 5' end by ligating a *Bgl* II-*Sal* I fragment from pDJ13 into pDJ11 that had been digested with *Bgl* II and *Sal* I. The resultant plasmid (pDJ14) was then digested with *Xba* I and *Sal* I and the entire AFM cDNA was cloned into the mammalian expression vector pDSRa (European Patent Application A20398753) that was modified to include unique *Xba* I and *Sal* I sites. The AFM expression vector was used to transfect a Chinese Hamster Ovary (CHO) cell line deficient in dihydrofolate reductase (CHO D<sup>-</sup>). Transfectants were selected in medium lacking hypoxanthine and thymidine. An RNase protection assay was used to screen for transfectants that had AFM-specific mRNA. A single clone was grown without serum as described (See, Bourdrel, L. et al., *Protein Expression*

Purif. 4: 130 (1993) to generate conditioned medium (CM) containing rAFM.

A single stable transfectant expressing high levels of AFM-specific mRNA was isolated. Immunoblots performed with AM339 (see Example 12) demonstrates that this antibody recognizes both rAFM produced from the transfectant as well as natural AFM isolated from human plasma (FIG. 5). rAFM was purified from CM derived from this transfectant (FIG. 6) and SDS-PAGE demonstrated rAFM to be greater than 95% pure with the same N-terminus as AFM purified from plasma (data not shown).

#### EXAMPLE 8: Protein Purification

Serum-free CM was concentrated 10-fold and diafiltered against 25 mM sodium phosphate, pH 7.5 using a Filtron ultrafiltration apparatus loaded with 10 K molecular weight cut-off filters. The CM was then adjusted with 3.9 M ammonium sulfate to achieve a final ammonium sulfate concentration of 1.6 M. This mixture was stirred for 0.5 hr; no precipitate was observed. The solution was then filtered in succession through 0.45mm and 0.22mm filters. The resultant filtrate was loaded onto a Phenyl Sepharose HP column previously equilibrated with 50 mM sodium phosphate, 1.6 M ammonium sulfate; pH 7.5. Loading, washing and elution were done at a linear flow rate of 40 cm/hr. Following sample loading, the column was washed successively with 50 mM sodium phosphate, 1.6 M ammonium sulfate, pH 7.5 and 50 mM sodium phosphate, pH 7.5. rAFM was eluted with water and detected by immunoblotting (see below).

The pH of the water eluate was adjusted to 8.0 with 0.1 M Tris-base and loaded at a flow rate of 100 cm/hr onto a Q Sepharose HP column previously equilibrated with 20 mM Tris-HCl, pH 8.0. Following sample loading, the column was thoroughly washed with 20 mM Tris-HCl, pH 8.2, then with 0.1 M NaCl, 20 mM Tris-

HCl, pH 8.2. The column was then resolved with a 3 column volume (CV) linear gradient of NaCl ranging from 0 - 0.3 M, followed by a 1 CV linear gradient of NaCl ranging from 0.3 M to 1.0 M in 20 mM Tris-HCl, pH 8.0.

- 5 Fractions containing rAFM were pooled and loaded onto a Superdex 200 column previously equilibrated with phosphate-buffered saline pH 7.4. The column was loaded and resolved at a flow rate of 1 ml/min.

10 EXAMPLE 9: Delipidation of AFM

- To one volume of protein solution was added 2.5 volumes of 1-butano/diisopropyl ether (40:60). The mixture was shaken gently for 30 minutes at room temperature and then centrifuged for 5 minutes at 500 x g. The aqueous layer was recovered and to this was added a second 2.5 volumes of 1-butanol/diisopropyl ether (40:60). The sample was treated as before and the aqueous phase recovered. To the twice extracted aqueous phase was added one volume of diisopropyl ether, the mixture shaken and immediately centrifuged at 250 x g for 5 minutes. The aqueous layer was recovered and used as essentially delipidated.

EXAMPLE 10: Immunoblotting

- 25 All samples for immunoblotting were electrophoresed on 4-20% SDS polyacrylamide gels (Novex) and the proteins were then electrophoretically transferred to nitrocellulose membranes (Schleicher and Schuell, cat. no. BA83). The filters were first treated with polyclonal antibody AM339 that was raised in rabbits against a synthetic peptide (DLSLREGKFTDSENV) (SEQ ID NO: 30) derived from amino acids 304 - 319 in AFM and then with donkey anti-rabbit Ig linked to horseradish peroxidase. Immune complexes were visualized by enhanced chemiluminescence according to the manufacturer's (Amersham) specifications.

EXAMPLE 11: Nucleotide sequence comparisons

Nucleotide sequences were compared to the Genbank (Release 78.0) sequences using the method of Pearson and Lipman (*Proc. Natl. Acad. Sci. USA* **84**, 2444-2448 (1988)) in the FASTA program of the Genetics Computer Group (GCG), Inc. (Madison, WI). Comparisons are shown in FIG. 2A.

EXAMPLE 12: Generation of peptide antibodies specific for AFM

Since AFM was similar to at least three other members of the albumin family, it was possible that antibodies generated against the entire protein would also cross react with other albumin family proteins. To create AFM-specific antibodies, we first identified specific peptides of AFM which were dissimilar to peptide sequences found in the other albumin family proteins. These peptide sequences were then synthesized and the synthetic peptides were used to inject rabbits to obtain polyclonal antisera against each peptide. The AFM peptides that were synthesized are as follows:

H1    EKLVKDMVEYKDRC                    (SEQ ID NO:31)  
      aa 43-56 in AFM

25    The corresponding antibody is referred to as AM384.

H2    CIINSNKDDRPKDLSLR                (SEQ ID NO:32)  
      aa 292-308 in AFM

The corresponding antibody is referred to as AM 609.

30

H3    DLSLREGKFTDSENV ·                (SEQ ID NO:30)  
      aa 304-319 in AFM

The corresponding antibody is referred to as AM 339.

H4 CQERDADPDTFFAKFT  
aa 319-334 in AFM

(SEQ ID NO:33)

The corresponding antibody is referred to as AM 1104.

5 EXAMPLE 13: Generation of polyclonal antisera to AFM

Rabbit polyclonal antiserum was generated from purified AFM by methods that are standard in the art. The antiserum was found to bind specifically to AFM.

10 Discussion of Examples

ALB family proteins are comprised of three homologous folding domains and are predicted to have evolved from an ancestral gene that coded for an approximately 190 amino acid single domain protein  
15 containing 3 double loops formed by 6 disulfide bridges. The genes in this family have all been mapped to the 4q11-q22 region of chromosome 4. AFM shares significant homology with ALB family proteins and has Cys residues consistent with a similar overall 3-domain organization.  
20 In addition, AFM has been localized to chromosome 4. Thus, there is compelling evidence that AFM is the fourth member of the ALB family.

There are some noteworthy distinctions among ALB family members. Concentrations in adult serum are 50  
25 ng/ml for AFP, 350 mg/ml for VDB, 40 mg/ml for ALB and 30 mg/ml for AFM (data obtained by immunoblot analysis). ALB is not glycosylated, AFP and VDB each have 1 potential N-glycosylation site while AFM has 4 potential sites. ALB expresses one free thiol group that has been  
30 implicated in complex formation with Cys, glutathione, mercurial and gold compounds. In contrast, AFP and VDB have an even number of Cys residues and are thought not to have a free thiol. AFM has an even number of Cys residues, suggesting that it may not have a free thiol  
35 and may not bind glutathione and mercurials as does ALB.

There also are differences in the intradomain disulfide bonding pattern among ALB family members. VDB is predicted to have a disulfide bridge in double loop 1A. This bridge is absent in ALB, AFP and AFM. A  
5 disulfide bridge domain 2C is common to ALB, VDB and AFM but is not present in AFP. Thus, while the 4 ALB family proteins are evolutionarily related, there are clear differences in the molecular organization of these proteins.

10 Structural similarities between AFM and other ALB family members suggest that AFM could scavenge or transport a variety of ligands. We examined whether known ligand-binding sites in the sequence of ALB family  
15 proteins were also present in AFM. VDB has a binding site for sterols between amino acids 35-49 and a binding site for actin between amino acids 373-403. Using the GCG GAP program and the alignment of FIG. 2A, AFM has 60% similarity and 40% identity between VDB amino acids 35-49 but only 32% similarity and 10% identity between  
20 VDB amino acids 373-403. Thus, it is possible that AFM has sterol binding sites (e.g., the amino acids in AFM that correspond to amino acids 35 - 49 of VDB) but it is not likely to bind actin. The X-ray crystal structure of ALB was used to show that ALB binds a variety of  
25 ligands (aspirin, warfarin, IS, DIS, TIB, bilirubin) between amino acids 186-260 in domain 2A and an array of ligands (aspirin, diazepam, digitoxin, clofibrate, ibuprofen, IS, DIS, TIB, long chain fatty acids) between amino acids 379-455 in domain 3A. A GCG GAP comparison  
30 between analogous regions in AFM reveals that AFM has 54% similarity and 35% identity in domain 2A and 45% similarity and 25% identity in domain 3A. These moderate degrees of similarity make it possible but not conclusive as to whether AFM binds the same ligands as  
35 ALB in domains 2A and 3A.



Abbreviations

The abbreviations used in the above Examples  
5 are: AFM, afamin; AFP, a-fetoprotein; ALB, human serum  
albumin; CHO, Chinese Hamster Ovary; CM, conditioned  
medium; CV, column volume; DIS, 3, 5- diiodosalicylic  
acid; DTT, dithiothreitol; HPLC, high performance liquid  
chromotography; IS, 5-iodosalicylic acid; PAGE,  
10 polyacrylamide gel electrophoresis; PBS, phosphate-  
buffered saline; PCR, polymerase chain reaction; r,  
recombinant; TIB, 2,3,5- triiodobenzoic acid; VDP,  
vitamin D-binding protein.

15

\* \* \*

The invention now being fully described, it  
will be apparent to one of ordinary skill in the art  
that many changes and modifications can be made thereto,  
20 without departing from the spirit and scope of the  
invention as set forth herein.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: AMGEN INC.
- (ii) TITLE OF INVENTION: Afamin: A Human Serum Albumin-Like Protein
- (iii) NUMBER OF SEQUENCES: 33
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Amgen Center, Patent Operations/RRC
  - (B) STREET: 1840 DeHavilland Drive
  - (C) CITY: Thousand Oaks
  - (D) STATE: California
  - (E) COUNTRY: U.S.
  - (F) ZIP: 91320-1789
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2287 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 318..2117
- (ix) FEATURE:
  - (A) NAME/KEY: mat\_peptide
  - (B) LOCATION: 381..2114
- (ix) FEATURE:
  - (A) NAME/KEY: sig\_peptide
  - (B) LOCATION: 318..380

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCCGAGTCT CTGCGCCTTC ACATAGTTGT CACAGGACTA AAGCAAATTG ATCCAGGGGG	60
AAACACTGTA GACCGTGTAT ATAAAAACAC TCTATAAACT GCAATGCTCA ATTCTTAGTA	120
TAACTATTGT TGTGTATTG ATATTTATTA GTATTGGTGC TCACAAAAAG AGTCTAAATT	180
CCATAAGTCT TTATATTCAG GCTACTCTTT ATTTTGGAAA ACTCATTTTC TATCACCTTT	240
TTCTATTTTA CTCCATATTG AGGCCTCATA AATCCAATTT TTTATTTCTT TCTTTTGTA	300
ATGTGGTTTC TACAAAG ATG AAA CTA CTA AAA CTT ACA GGT TTT ATT TTT	350
Met Lys Leu Leu Lys Leu Thr Gly Phe Ile Phe	
-21 -20 -15	
TTC TTG TTT TTT TTG ACT GAA TCC CTA ACC CTG CCC ACA CAA CCT CGG	398
Phe Leu Phe Phe Leu Thr Glu Ser Leu Thr Leu Pro Thr Gln Pro Arg	
-10 -5 1 5	
GAT ATA GAG AAC TTC AAT AGT ACT CAA AAA TTT ATA GAA GAT AAT ATT	446
Asp Ile Glu Asn Phe Asn Ser Thr Gln Lys Phe Ile Glu Asp Asn Ile	
10 15 20	
GAA TAC ATC ACC ATC ATT GCA TTT GCT CAG TAT GTT CAG GAA GCA ACC	494
Glu Tyr Ile Thr Ile Ile Ala Phe Ala Gln Tyr Val Gln Glu Ala Thr	
25 30 35	
TTT GAA GAA ATG GAA AAG CTG GTG AAA GAC ATG GTA GAA TAC AAA GAC	542
Phe Glu Glu Met Glu Lys Leu Val Lys Asp Met Val Glu Tyr Lys Asp	
40 45 50	
AGA TGT ATG GCT GAC AAG ACG CTC CCA GAG TGT TCA AAA TTA CCT AAT	590
Arg Cys Met Ala Asp Lys Thr Leu Pro Glu Cys Ser Lys Leu Pro Asn	
55 60 65 70	
AAT GTT TTA CAG GAA AAA ATA TGT GCT ATG GAG GGG CTG CCA CAA AAG	638
Asn Val Leu Gln Glu Lys Ile Cys Ala Met Glu Gly Leu Pro Gln Lys	
75 80 85	
CAT AAT TTC TCA CAC TGC TGC AGT AAG GTT GAT GCT CAA AGA AGA CTC	686
His Asn Phe Ser His Cys Cys Ser Lys Val Asp Ala Gln Arg Arg Leu	
90 95 100	
TGT TTC TTC TAT AAC AAG AAA TCT GAT GTG GGA TTT CTG CCT CCT TTC	734
Cys Phe Phe Tyr Asn Lys Lys Ser Asp Val Gly Phe Leu Pro Pro Phe	
105 110 115	
CCT ACC CTG GAT CCC GAA GAG AAA TGC CAG GCT TAT GAA AGT AAC AGA	782
Pro Thr Leu Asp Pro Glu Glu Lys Cys Gln Ala Tyr Glu Ser Asn Arg	
120 125 130	
GAA TCC CTT TTA AAT CAC TTT TTA TAT GAA GTT GCC AGA AGG AAC CCA	830
Glu Ser Leu Leu Asn His Phe Leu Tyr Glu Val Ala Arg Arg Asn Pro	
135 140 145 150	
TTT GTC TTC GCC CCT ACA CTT CTA ACT GTT GCT GTT CAT TTT GAG GAG	878

Phe	Val	Phe	Ala	Pro	Thr	Leu	Leu	Thr	Val	Ala	Val	His	Phe	Glu	Glu	
				155					160					165		
GTG	GCC	AAA	TCA	TGT	TGT	GAA	GAA	CAA	AAC	AAA	GTC	AAC	TGC	CTT	CAA	926
Val	Ala	Lys	Ser	Cys	Cys	Glu	Glu	Gln	Asn	Lys	Val	Asn	Cys	Leu	Gln	
			170					175					180			
ACA	AGG	GCA	ATA	CCT	GTC	ACA	CAA	TAT	TTA	AAA	GCA	TTT	TCT	TCT	TAT	974
Thr	Arg	Ala	Ile	Pro	Val	Thr	Gln	Tyr	Leu	Lys	Ala	Phe	Ser	Ser	Tyr	
		185					190					195				
CAA	AAA	CAT	GTC	TGT	GGG	GCA	CTT	TTG	AAA	TTT	GGA	ACC	AAA	GTT	GTA	1022
Gln	Lys	His	Val	Cys	Gly	Ala	Leu	Leu	Lys	Phe	Gly	Thr	Lys	Val	Val	
	200					205					210					
CAC	TTT	ATA	TAT	ATT	GCG	ATA	CTC	AGT	CAA	AAA	TTC	CCC	AAG	ATT	GAA	1070
His	Phe	Ile	Tyr	Ile	Ala	Ile	Leu	Ser	Gln	Lys	Phe	Pro	Lys	Ile	Glu	
215					220					225					230	
TTT	AAG	GAG	CTT	ATT	TCT	CTT	GTA	GAA	GAT	GTT	TCT	TCC	AAC	TAT	GAT	1118
Phe	Lys	Glu	Leu	Ile	Ser	Leu	Val	Glu	Asp	Val	Ser	Ser	Asn	Tyr	Asp	
			235					240						245		
GGA	TGC	TGT	GAA	GGG	GAT	GTT	GTG	CAG	TGC	ATC	CGT	GAC	ACG	AGC	AAG	1166
Gly	Cys	Cys	Glu	Gly	Asp	Val	Val	Gln	Cys	Ile	Arg	Asp	Thr	Ser	Lys	
		250						255					260			
GTT	ATG	AAC	CAT	ATT	TGT	TCA	AAA	CAA	GAT	TCT	ATC	TCC	AGC	AAA	ATC	1214
Val	Met	Asn	His	Ile	Cys	Ser	Lys	Gln	Asp	Ser	Ile	Ser	Ser	Lys	Ile	
		265					270					275				
AAA	GAG	TGC	TGT	GAA	AAG	AAA	ATA	CCA	GAG	CGC	GGC	CAG	TGC	ATA	ATT	1262
Lys	Glu	Cys	Cys	Glu	Lys	Lys	Ile	Pro	Glu	Arg	Gly	Gln	Cys	Ile	Ile	
	280					285					290					
AAC	TCA	AAC	AAA	GAT	GAT	AGA	CCA	AAG	GAT	TTA	TCT	CTA	AGA	GAA	GGA	1310
Asn	Ser	Asn	Lys	Asp	Asp	Arg	Pro	Lys	Asp	Leu	Ser	Leu	Arg	Glu	Gly	
295				300					305					310		
AAA	TTT	ACT	GAC	AGT	GAA	AAT	GTG	TGT	CAA	GAA	CGA	GAT	GCT	GAC	CCA	1358
Lys	Phe	Thr	Asp	Ser	Glu	Asn	Val	Cys	Gln	Glu	Arg	Asp	Ala	Asp	Pro	
			315					320					325			
GAC	ACC	TTC	TTT	GCG	AAG	TTT	ACT	TTT	GAA	TAC	TCA	AGG	AGA	CAT	CCA	1406
Asp	Thr	Phe	Phe	Ala	Lys	Phe	Thr	Phe	Glu	Tyr	Ser	Arg	Arg	His	Pro	
		330					335						340			
GAC	CTG	TCT	ATA	CCA	GAG	CTT	TTA	AGA	ATT	GTT	CAA	ATA	TAC	AAA	GAT	1454
Asp	Leu	Ser	Ile	Pro	Glu	Leu	Arg	Ile	Val	Gln	Ile	Tyr	Lys	Asp		
		345				350					355					
CTC	CTG	AGA	AAT	TGC	TGC	AAC	ACA	GAA	AAC	CCT	CCA	GGT	TGT	TAC	CGT	1502
Leu	Leu	Arg	Asn	Cys	Cys	Asn	Thr	Glu	Asn	Pro	Pro	Gly	Cys	Tyr	Arg	
	360					365				370						
TAC	GCG	GAA	GAC	AAA	TTC	AAT	GAG	ACA	ACT	GAG	AAA	AGC	CTC	AAG	ATG	1550
Tyr	Ala	Glu	Asp	Lys	Phe	Asn	Glu	Thr	Thr	Glu	Lys	Ser	Leu	Lys	Met	
375				380					385					390		

GTA CAA CAA GAA TGT AAA CAT TTC CAG AAT TTG GGG AAG GAT GGT TTG Val Gln Gln Glu Cys Lys His Phe Gln Asn Leu Gly Lys Asp Gly Leu 395 400 405	1598
AAA TAC CAT TAC CTC ATC AGG CTC ACG AAG ATA GCT CCC CAA CTC TCC Lys Tyr His Tyr Leu Ile Arg Leu Thr Lys Ile Ala Pro Gln Leu Ser 410 415 420	1646
ACT GAA GAA CTG GTG TCT CTT GGC GAG AAA ATG GTG ACA GCT TTC ACT Thr Glu Glu Leu Val Ser Leu Gly Glu Lys Met Val Thr Ala Phe Thr 425 430 435	1694
ACT TGC TGT ACG CTA AGT GAA GAG TTT GCC TGT GTT GAT AAT TTG GCA Thr Cys Cys Thr Leu Ser Glu Glu Phe Ala Cys Val Asp Asn Leu Ala 440 445 450	1742
GAT TTA GTT TTT GGA GAG TTA TGT GGA GTA AAT GAA AAT CGA ACT ATC Asp Leu Val Phe Gly Glu Leu Cys Gly Val Asn Glu Asn Arg Thr Ile 455 460 465 470	1790
AAC CCT GCT GTG GAC CAC TGC TGT AAA ACA AAC TTT GCC TTC AGA AGG Asn Pro Ala Val Asp His Cys Cys Lys Thr Asn Phe Ala Phe Arg Arg 475 480 485	1838
CCC TGC TTT GAG AGT TTG AAA GCT GAT AAA ACA TAT GTG CCT CCA CCT Pro Cys Phe Glu Ser Leu Lys Ala Asp Lys Thr Tyr Val Pro Pro Pro 490 495 500	1886
TTC TCT CAA GAT TTA TTT ACC TTT CAC GCA GAC ATG TGT CAA TCT CAG Phe Ser Gln Asp Leu Phe Thr Phe His Ala Asp Met Cys Gln Ser Gln 505 510 515	1934
AAT GAG GAG CTT CAG AGG AAG ACA GAC AGG TTT CTT GTC AAC TTA GTG Asn Glu Glu Leu Gln Arg Lys Thr Asp Arg Phe Leu Val Asn Leu Val 520 525 530	1982
AAG CTG AAG CAT GAA CTC ACA GAT GAA GAG CTG CAG TCT TTG TTT ACA Lys Leu Lys His Glu Leu Thr Asp Glu Glu Leu Gln Ser Leu Phe Thr 535 540 545 550	2030
AAT TTC GCA AAT GTA GTG GAT AAG TGC TGC AAA GCA GAG AGT CCT GAA Asn Phe Ala Asn Val Val Asp Lys Cys Cys Lys Ala Glu Ser Pro Glu 555 560 565	2078
GTC TGC TTT AAT GAA GAG AGT CCA AAA ATT GGC AAC TGAAGCCAGC Val Cys Phe Asn Glu Glu Ser Pro Lys Ile Gly Asn 570 575	2124
TGCTGGAGAT ATGTAAAGAA AAAAGCACCA AAGGGAAGGC TTCCTATCTG TGTGGTGATG	2184
AATCGCATTT CCTGAGAACA AAATAAAAGG ATTTTCTGT AACTGTCACC TGAAATAATA	2244
CATTGCAGCA AGCAATAAAC ACAACATTTT GTAAAGTTAA AAA	2287

(2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 599 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Leu Leu Lys Leu Thr Gly Phe Ile Phe Phe Leu Phe Phe Leu
-21 -20                      -15                      -10

Thr Glu Ser Leu Thr Leu Pro Thr Gln Pro Arg Asp Ile Glu Asn Phe
-5                      1                      5                      10

Asn Ser Thr Gln Lys Phe Ile Glu Asp Asn Ile Glu Tyr Ile Thr Ile
15                      20                      25

Ile Ala Phe Ala Gln Tyr Val Gln Glu Ala Thr Phe Glu Glu Met Glu
30                      35                      40

Lys Leu Val Lys Asp Met Val Glu Tyr Lys Asp Arg Cys Met Ala Asp
45                      50                      55

Lys Thr Leu Pro Glu Cys Ser Lys Leu Pro Asn Asn Val Leu Gln Glu
60                      65                      70                      75

Lys Ile Cys Ala Met Glu Gly Leu Pro Gln Lys His Asn Phe Ser His
80                      85                      90

Cys Cys Ser Lys Val Asp Ala Gln Arg Arg Leu Cys Phe Phe Tyr Asn
95                      100                      105

Lys Lys Ser Asp Val Gly Phe Leu Pro Pro Phe Pro Thr Leu Asp Pro
110                      115                      120

Glu Glu Lys Cys Gln Ala Tyr Glu Ser Asn Arg Glu Ser Leu Leu Asn
125                      130                      135

His Phe Leu Tyr Glu Val Ala Arg Arg Asn Pro Phe Val Phe Ala Pro
140                      145                      150                      155

Thr Leu Leu Thr Val Ala Val His Phe Glu Glu Val Ala Lys Ser Cys
160                      165                      170

Cys Glu Glu Gln Asn Lys Val Asn Cys Leu Gln Thr Arg Ala Ile Pro
175                      180                      185

Val Thr Gln Tyr Leu Lys Ala Phe Ser Ser Tyr Gln Lys His Val Cys
190                      195                      200

Gly Ala Leu Leu Lys Phe Gly Thr Lys Val Val His Phe Ile Tyr Ile
205                      210                      215

Ala Ile Leu Ser Gln Lys Phe Pro Lys Ile Glu Phe Lys Glu Leu Ile
220                      225                      230                      235

Ser Leu Val Glu Asp Val Ser Ser Asn Tyr Asp Gly Cys Cys Glu Gly

```

240										245					250				
Asp	Val	Val	Gln	Cys	Ile	Arg	Asp	Thr	Ser	Lys	Val	Met	Asn	His	Ile				
			255					260					265						
Cys	Ser	Lys	Gln	Asp	Ser	Ile	Ser	Ser	Lys	Ile	Lys	Glu	Cys	Cys	Glu				
		270					275					280							
Lys	Lys	Ile	Pro	Glu	Arg	Gly	Gln	Cys	Ile	Ile	Asn	Ser	Asn	Lys	Asp				
	285					290					295								
Asp	Arg	Pro	Lys	Asp	Leu	Ser	Leu	Arg	Glu	Gly	Lys	Phe	Thr	Asp	Ser				
300					305					310					315				
Glu	Asn	Val	Cys	Gln	Glu	Arg	Asp	Ala	Asp	Pro	Asp	Thr	Phe	Phe	Ala				
				320					325						330				
Lys	Phe	Thr	Phe	Glu	Tyr	Ser	Arg	Arg	His	Pro	Asp	Leu	Ser	Ile	Pro				
			335					340					345						
Glu	Leu	Leu	Arg	Ile	Val	Gln	Ile	Tyr	Lys	Asp	Leu	Leu	Arg	Asn	Cys				
	350						355					360							
Cys	Asn	Thr	Glu	Asn	Pro	Pro	Gly	Cys	Tyr	Arg	Tyr	Ala	Glu	Asp	Lys				
	365					370					375								
Phe	Asn	Glu	Thr	Thr	Glu	Lys	Ser	Leu	Lys	Met	Val	Gln	Gln	Glu	Cys				
380					385					390					395				
Lys	His	Phe	Gln	Asn	Leu	Gly	Lys	Asp	Gly	Leu	Lys	Tyr	His	Tyr	Leu				
				400					405				410						
Ile	Arg	Leu	Thr	Lys	Ile	Ala	Pro	Gln	Leu	Ser	Thr	Glu	Glu	Leu	Val				
			415					420					425						
Ser	Leu	Gly	Glu	Lys	Met	Val	Thr	Ala	Phe	Thr	Thr	Cys	Cys	Thr	Leu				
		430					435					440							
Ser	Glu	Glu	Phe	Ala	Cys	Val	Asp	Asn	Leu	Ala	Asp	Leu	Val	Phe	Gly				
	445					450					455								
Glu	Leu	Cys	Gly	Val	Asn	Glu	Asn	Arg	Thr	Ile	Asn	Pro	Ala	Val	Asp				
460					465					470					475				
His	Cys	Cys	Lys	Thr	Asn	Phe	Ala	Phe	Arg	Arg	Pro	Cys	Phe	Glu	Ser				
				480					485					490					
Leu	Lys	Ala	Asp	Lys	Thr	Tyr	Val	Pro	Pro	Pro	Phe	Ser	Gln	Asp	Leu				
			495					500					505						
Phe	Thr	Phe	His	Ala	Asp	Met	Cys	Gln	Ser	Gln	Asn	Glu	Glu	Leu	Gln				
	510					515						520							
Arg	Lys	Thr	Asp	Arg	Phe	Leu	Val	Asn	Leu	Val	Lys	Leu	Lys	His	Glu				
	525					530					535								
Leu	Thr	Asp	Glu	Glu	Leu	Gln	Ser	Leu	Phe	Thr	Asn	Phe	Ala	Asn	Val				
540					545					550					555				

Val Asp Lys Cys Cys Lys Ala Glu Ser Pro Glu Val Cys Phe Asn Glu  
560 565 570

Glu Ser Pro Lys Ile Gly Asn  
575

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 609 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala  
1 5 10 15

Tyr Ser Arg Gly Val Phe Arg Arg Asp Ala His Lys Ser Glu Val Ala  
20 25 30

His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu  
35 40 45

Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val  
50 55 60

Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp  
65 70 75 80

Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp  
85 90 95

Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala  
100 105 110

Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln  
115 120 125

His	Lys	Asp	Asp	Asn	Pro	Asn	Leu	Pro	Arg	Leu	Val	Arg	Pro	Glu	Val
	130					135					140				

Asp	Val	Met	Cys	Thr	Ala	Phe	His	Asp	Asn	Glu	Glu	Thr	Phe	Leu	Lys
145					150					155					160

Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro  
165 170 175

Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys  
180 185 190



Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu  
 195 200 205  
 Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys  
 210 215 220  
 Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val  
 225 230 235 240  
 Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser  
 245 250 255  
 Lys Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly  
 260 265 270  
 Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile  
 275 280 285  
 Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu  
 290 295 300  
 Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp  
 305 310 315 320  
 Glu Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser  
 325 330 335  
 Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly  
 340 345 350  
 Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val  
 355 360 365  
 Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys  
 370 375 380  
 Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu  
 385 390 395 400  
 Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys  
 405 410 415  
 Glu Leu Phe Lys Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu  
 420 425 430  
 Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val  
 435 440 445  
 Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His  
 450 455 460  
 Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val  
 465 470 475 480  
 Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg  
 485 490 495  
 Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe

500	505	510
Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala		
515	520	525
Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu		
530	535	540
Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys		
545	550	555
Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala		
565	570	575
Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe		
580	585	590
Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly		
595	600	605
Leu		

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 609 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Lys Trp Val Glu Ser Ile Phe Leu Ile Phe Leu Leu Asn Phe Thr		
1	5	10
Glu Ser Arg Thr Leu His Arg Asn Glu Tyr Gly Ile Ala Ser Ile Leu		
20	25	30
Asp Ser Tyr Gln Cys Thr Ala Glu Ile Ser Leu Ala Asp Leu Ala Thr		
35	40	45
Ile Phe Phe Ala Gln Phe Val Gln Glu Ala Thr Tyr Lys Glu Val Ser		
50	55	60
Lys Met Val Lys Asp Ala Leu Thr Ala Ile Glu Lys Pro Thr Gly Asp		
65	70	75
Glu Gln Ser Ser Gly Cys Leu Glu Asn Gln Leu Pro Ala Phe Leu Glu		
85	90	95
Glu Leu Cys His Glu Lys Glu Ile Leu Glu Lys Tyr Gly His Ser Asp		
100	105	110

Cys Cys Ser Gln Ser Glu Glu Gly Arg His Asn Cys Phe Leu Ala His  
 115 120 125  
 Lys Lys Pro Thr Pro Ala Ser Ile Pro Leu Phe Gln Val Pro Glu Pro  
 130 135 140  
 Val Thr Ser Cys Glu Ala Tyr Glu Glu Asp Arg Glu Thr Phe Met Asn  
 145 150 155 160  
 Lys Phe Ile Tyr Glu Ile Ala Arg Arg His Pro Phe Leu Tyr Ala Pro  
 165 170 175  
 Thr Ile Leu Leu Trp Ala Ala Arg Tyr Asp Lys Ile Ile Pro Ser Cys  
 180 185 190  
 Cys Lys Ala Glu Asn Ala Val Glu Cys Phe Gln Thr Lys Ala Ala Thr  
 195 200 205  
 Val Thr Lys Glu Leu Arg Glu Ser Ser Leu Leu Asn Gln His Ala Cys  
 210 215 220  
 Ala Val Met Lys Asn Phe Gly Thr Arg Thr Phe Gln Ala Ile Thr Val  
 225 230 235 240  
 Thr Lys Leu Ser Gln Lys Phe Thr Lys Val Asn Phe Thr Glu Ile Gln  
 245 250 255  
 Lys Leu Val Leu Asp Val Ala His Val His Glu His Cys Cys Arg Gly  
 260 265 270  
 Asp Val Leu Asp Cys Leu Gln Asp Gly Glu Lys Ile Met Ser Tyr Ile  
 275 280 285  
 Cys Ser Gln Gln Asp Thr Leu Ser Asn Lys Ile Thr Glu Cys Cys Lys  
 290 295 300  
 Leu Thr Thr Leu Glu Arg Gly Gln Cys Ile Ile His Ala Glu Asn Asp  
 305 310 315 320  
 Glu Lys Pro Glu Gly Leu Ser Pro Asn Leu Asn Arg Phe Leu Gly Asp  
 325 330 335  
 Arg Asp Phe Asn Gln Phe Ser Ser Gly Glu Lys Asn Ile Phe Leu Ala  
 340 345 350  
 Ser Phe Val His Glu Tyr Ser Arg Arg His Pro Gln Leu Ala Val Ser  
 355 360 365  
 Val Ile Leu Arg Val Ala Lys Gly Tyr Gln Glu Leu Leu Glu Lys Cys  
 370 375 380  
 Phe Gln Thr Glu Asn Pro Leu Glu Cys Gln Asp Lys Gly Glu Glu Glu  
 385 390 395 400  
 Leu Gln Lys Tyr Ile Gln Glu Ser Gln Ala Leu Ala Lys Arg Ser Cys  
 405 410 415  
 Gly Leu Phe Gln Lys Leu Gly Glu Tyr Tyr Leu Gln Asn Ala Phe Leu

				420						425						430	
Val	Ala	Tyr	Thr	Lys	Lys	Ala	Pro	Gln	Leu	Thr	Ser	Ser	Glu	Leu	Met		
435				440				445									
Ala	Ile	Thr	Arg	Lys	Met	Ala	Ala	Thr	Ala	Ala	Thr	Cys	Cys	Gln	Leu		
450				455				460									
Ser	Glu	Asp	Lys	Leu	Leu	Ala	Cys	Gly	Glu	Gly	Ala	Ala	Asp	Ile	Ile		
465				470				475				480					
Ile	Gly	His	Leu	Cys	Ile	Arg	His	Glu	Met	Thr	Pro	Val	Asn	Pro	Gly		
				485				490				495					
Val	Gly	Gln	Cys	Cys	Thr	Ser	Ser	Tyr	Ala	Asn	Arg	Arg	Pro	Cys	Phe		
				500				505				510					
Ser	Ser	Leu	Val	Val	Asp	Glu	Thr	Tyr	Val	Pro	Pro	Ala	Phe	Ser	Asp		
515				520				525									
Asp	Lys	Phe	Ile	Phe	His	Lys	Asp	Leu	Cys	Gln	Ala	Gln	Gly	Val	Ala		
530				535				540									
Leu	Gln	Thr	Met	Lys	Gln	Glu	Phe	Leu	Ile	Asn	Leu	Val	Lys	Gln	Lys		
545				550				555				560					
Pro	Gln	Ile	Thr	Glu	Glu	Gln	Leu	Glu	Ala	Val	Ile	Ala	Asp	Phe	Ser		
				565				570				575					
Gly	Leu	Leu	Glu	Lys	Cys	Cys	Gln	Gly	Gln	Glu	Gln	Glu	Val	Cys	Phe		
580				585				590									
Ala	Glu	Glu	Gly	Gln	Lys	Leu	Ile	Ser	Lys	Thr	Arg	Ala	Ala	Leu	Gly		
595				600				605									
Val																	

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 474 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Lys	Arg	Val	Leu	Val	Leu	Leu	Leu	Ala	Val	Ala	Phe	Gly	His	Ala
1				5					10					15	
Leu	Glu	Arg	Gly	Arg	Asp	Tyr	Glu	Lys	Asn	Lys	Val	Cys	Lys	Glu	Phe
			20					25					30		

Ser His Leu Gly Lys Glu Asp Phe Thr Ser Leu Ser Leu Val Leu Tyr  
 35 40 45  
 Ser Arg Lys Phe Pro Ser Gly Thr Phe Glu Gln Val Ser Gln Leu Val  
 50 55 60  
 Lys Glu Val Val Ser Leu Thr Glu Ala Cys Cys Ala Glu Gly Ala Asp  
 65 70 75 80  
 Pro Asp Cys Tyr Asp Thr Arg Thr Ser Ala Leu Ser Ala Lys Ser Cys  
 85 90 95  
 Glu Ser Asn Ser Pro Phe Pro Val His Pro Gly Thr Ala Glu Cys Cys  
 100 105 110  
 Thr Lys Glu Gly Leu Glu Arg Lys Leu Cys Met Ala Ala Leu Lys His  
 115 120 125  
 Gln Pro Gln Glu Phe Pro Thr Tyr Val Glu Pro Thr Asn Asp Glu Ile  
 130 135 140  
 Cys Glu Ala Phe Arg Lys Asp Pro Lys Glu Tyr Ala Asn Gln Phe Met  
 145 150 155 160  
 Trp Glu Tyr Ser Thr Asn Tyr Gly Gln Ala Pro Leu Ser Leu Leu Val  
 165 170 175  
 Ser Tyr Thr Lys Ser Tyr Leu Ser Met Val Gly Ser Cys Cys Thr Ser  
 180 185 190  
 Ala Ser Pro Thr Val Cys Phe Leu Lys Glu Arg Leu Gln Leu Lys His  
 195 200 205  
 Leu Ser Leu Leu Thr Thr Leu Ser Asn Arg Val Cys Ser Gln Tyr Ala  
 210 215 220  
 Ala Tyr Gly Glu Lys Lys Ser Arg Leu Ser Asn Leu Ile Lys Leu Ala  
 225 230 235 240  
 Gln Lys Val Pro Thr Ala Asp Leu Glu Asp Val Leu Pro Leu Ala Glu  
 245 250 255  
 Asp Ile Thr Asn Ile Leu Ser Lys Cys Cys Glu Ser Ala Ser Glu Asp  
 260 265 270  
 Cys Met Ala Lys Glu Leu Pro Glu His Thr Val Lys Leu Cys Asp Asn  
 275 280 285  
 Leu Ser Thr Lys Asn Ser Lys Phe Glu Asp Cys Cys Gln Glu Lys Thr  
 290 295 300  
 Ala Met Asp Val Phe Val Cys Thr Tyr Phe Met Pro Ala Ala Gln Leu  
 305 310 315 320  
 Pro Glu Leu Pro Asp Val Glu Leu Pro Thr Asn Lys Asp Val Cys Asp  
 325 330 335  
 Pro Gly Asn Thr Lys Val Met Asp Lys Tyr Thr Phe Glu Leu Ser Arg

340				345				350							
Arg	Thr	His	Leu	Pro	Glu	Val	Phe	Leu	Ser	Lys	Val	Leu	Glu	Pro	Thr
		355					360					365			
Leu	Lys	Ser	Leu	Gly	Glu	Cys	Cys	Asp	Val	Glu	Asp	Ser	Thr	Thr	Cys
	370					375					380				
Phe	Asn	Ala	Lys	Gly	Pro	Leu	Leu	Lys	Lys	Glu	Leu	Ser	Ser	Phe	Ile
385					390					395					400
Asp	Lys	Gly	Gln	Glu	Leu	Cys	Ala	Asp	Tyr	Ser	Glu	Asn	Thr	Phe	Thr
				405					410					415	
Glu	Tyr	Lys	Lys	Lys	Leu	Ala	Glu	Arg	Leu	Lys	Ala	Lys	Leu	Pro	Asp
			420					425					430		
Ala	Thr	Pro	Lys	Glu	Leu	Ala	Lys	Leu	Val	Asn	Lys	Arg	Ser	Asp	Phe
		435					440					445			
Ala	Ser	Asn	Cys	Cys	Ser	Ile	Asn	Ser	Pro	Pro	Leu	Tyr	Cys	Asp	Ser
	450					455					460				
Glu	Ile	Asp	Ala	Glu	Leu	Lys	Asn	Ile	Leu						
465						470									

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Leu Pro Thr Gln Pro Arg Asp Ile Glu Asn Phe Xaa Ser Thr Gln Lys
1          5          10          15
Phe Ile Glu Asp Asn Ile Glu Tyr Ile Thr Ile Ile Ala Phe Ala Gln
          20          25          30
Tyr Val Gln
          35

```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Tyr His Tyr Leu Ile Arg  
1 5

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Phe Thr Phe Glu Tyr Ser Arg  
1 5

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Phe Thr Asp Ser Glu Asn Val Cys Gln Glu Arg Asp Ala Asp Pro  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ile Val Gln Ile Tyr Lys Asp Leu Leu Arg  
1 5 10

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 15 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ile	Ala	Pro	Gln	Leu	Ser	Thr	Glu	Glu	Leu	Val	Ser	Leu	Gly	Glu
1				5					10					15

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 12 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Arg	His	Pro	Asp	Leu	Ser	Ile	Pro	Glu	Leu	Leu	Arg
1				5					10		

## (2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 13 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Glu	Ser	Leu	Leu	Asn	His	Phe	Leu	Tyr	Glu	Val	Ala	Arg
1				5					10			

## (2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown



(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Arg	Asn	Pro	Phe	Val	Phe	Ala	Pro	Thr	Leu	Leu	Thr	Val	Ala	Val	His
1				5				10					15		
Phe	Glu	Glu	Val	Ala	Lys	Ser	Cys	Cys							
			20				25								

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Gln	Lys	Phe	Ile	Glu	Asp	Asn
1				5		

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACGCTGAATT CGCCARAART TYATHGARGA YAA

33

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ile Val Gln Ile Tyr Lys Asp  
1 5

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ACGCTAAGCT TGCRTCYTTR TADATYTGNA CDAT

34

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Asp Asn Ile Glu Tyr Ile Thr  
1 5

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ACGCTGAATT CGCGAYAAYA THGARTAYAT HAC

33

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Phe Thr Phe Glu Tyr Ser  
1 5

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 31 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ACGCTAAGCT TGCNGARTAY TCRAANGTRA A

31

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TATGTGCTAT GGAGGGGC

18

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CAACCCTGCT GTGGACCA

18

(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 22 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCACATATGT TTTATCAGCT TT

22

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 38 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCACCTCTAG ACCACCATGA AACTACTAAA ACTTACAG

38

(2) INFORMATION FOR SEQ ID NO:27:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown
  - (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AATTTCTCAG GAGATCTTTG TATA

24

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

AAATATACAA AGATCTCCTG AGAA

24

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TCCCGGTCGA CTCAGTTGCC AATTTTGA C

31

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Asp	Leu	Ser	Leu	Arg	Glu	Gly	Lys	Phe	Thr	Asp	Ser	Glu	Asn	Val	Cys
1				5					10					15	

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: unknown

(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Glu Lys Leu Val Lys Asp Met Val Glu Tyr Lys Asp Arg Cys  
1 5 10

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Cys Ile Ile Asn Ser Asn Lys Asp Asp Arg Pro Lys Asp Leu Ser Leu  
1 5 10 15

Arg

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Cys Gln Glu Arg Asp Ala Asp Pro Asp Thr Phe Phe Ala Lys Phe Thr  
1 5 10 15

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

- 5           1. A purified and isolated polynucleotide encoding AFM polypeptide or a variant thereof possessing a biological activity specific to AFM.
- 10           2. The polynucleotide of claim 1 which is a DNA sequence.
3. The DNA sequence according to claim 2 which is a cDNA sequence or a biological replica thereof.
- 15           4. The cDNA sequence of claim 3, which encodes AFM.
5. The DNA sequence according to claim 3 which is a genomic DNA sequence or a biological replica thereof.
- 20           6. The DNA sequence of claim 3 further including an endogenous expression control DNA sequence.
7. The DNA sequence according to claim 3 which is a wholly or partially chemically synthesized DNA sequence or a biological replica thereof.
- 25           8. A DNA vector comprising a DNA sequence according to claim 3.
- 30           9. The vector of claim 8 wherein said DNA sequence is operatively linked to an expression control DNA sequence.

10. A host cell stably transformed or transfected with a DNA sequence according to claim 3 in a manner allowing the expression in said host cell of AFM polypeptide or a variant thereof possessing a biological activity specific to AFM.

11. A method for producing AFM polypeptide or a variant thereof possessing a biological activity specific to AFM, said method comprising growing a host cell according to claim 8 in a suitable nutrient medium and isolating AFM polypeptide or variant thereof from said cell or the growth medium.

12. Purified and isolated AFM polypeptide or a variant thereof possessing a biological activity specific to AFM.

13. The polypeptide of claim 12 comprising amino acid residues 1 through 578 of FIG 2.

14. The polypeptide of claim 12 comprising residues -23 through 578 of FIG 2.

15. Purified and isolated AFM polypeptide complexed with Apolipoprotein A1.

16. An antibody specific for AFM or a variant thereof.

17. A monoclonal antibody according to claim 16.



18. A humanized antibody according to claims 16 or 17.

5 19. A method for modulating a biological activity of AFM comprising contacting AFM with an antibody according to claim 16 or 17.

10 20. A method for detecting the capacity of a cell to synthesize AFM comprising hybridizing a detectable polynucleotide encoding AFM or a fragment thereof with RNA of said cell.

15 21. A method for detecting the capacity of a cell to synthesize AFM comprising reacting an antibody according to claim 16 with polypeptides produced by said cell.

20 22. A method for determining the presence of AFM in a sample body fluid, comprising contacting the body fluid with an antibody according to Claim 16.

23. An antisense polynucleotide for a polynucleotide encoding AFM.

25 24. An antisense polynucleotide for a DNA specifying an endogenous expression control DNA sequence of AFM.

30 25. A hybrid fusion polypeptide comprising, at its amino terminal, an AFM polypeptide or a variant thereof possessing a biological activity specific to AFM and, at its carboxy terminal, at least one constant domain of an immunoglobulin heavy chain or allelic variant thereof.

26. A polynucleotide encoding a hybrid fusion protein according to claim 25.

- 5                   27. A DNA sequence encoding a polypeptide having a biological property specific for AFM and selected from the group consisting of:
- 10                   (a) the DNA sequence set out in FIG 1;
- (b) a DNA which hybridizes under stringent conditions to the DNA of (a); and
- (c) a DNA sequence which, but for the redundancy of the genetic code, would hybridize under stringent conditions to a DNA sequence of (a) or (b).

1/12

## FIG. 1A

CCCCGAGTCTCTGCGCCTTCACATAGTTGTACAGGACTAAAGCAAATTGATCCAGGGGG	60
AAACACTGTAGACCGTGTATATAAAAACACTCTATAAACTGCAATGCTCAATTCTTAGTA	120
TAACTATTGTTGTTGTATTGATATTTATTAGTATTGGTGCTCACAAAAGAGTCTAAATT	180
CCATAAGTCTTTATATTCAGGCTACTCTTTATTTTGGAAACTCATTTTCTATCACCTTT	240
TTCTATTTTACTCCATATTGAGGCCTCATAAATCCAATTTTATTCTTTCTTTTGTA	300
m k l l k l t g f i f f l f f	-21
ATGTGGTTTCTACAAAGATGAAACTACTAAACTTACAGGTTTTATTTTTTTCTTGT	360
l t e s l t L P T Q P R D I E N F N S T	14
TTTTGACTGAATCCCTAACCCTGCCCACACAACCTCGGGATATAGAGAACTTCAATAGTA	420
Q K F I E D N I E Y I T I I A F A Q Y V	34
CTCAAAAATTTATAGAAGATAATATTGAATACATCACCATCATTGCATTGCTCAGTATG	480
Q E A T F E E M E K L V K D M V E Y K D	54
TTCAGGAAGCAACCTTTGAAGAAATGAAAAGCTGGTGAAAAGACATGGTAGAATACAAAG	540
R C M A D K T L P E C S K L P N N V L Q	74
ACAGATGTATGGCTGACAAGACGCTCCCAGAGTGTTCAAAATTACCTAATAATGTTTAC	600
E K I C A M E G L P Q K H N F S H C C S	94
AGGAAAAAATATGTGCTATGGAGGGGCTGCCACAAAAGCATAATTTCTCACACTGCTGCA	660
K V D A Q R R L C F F Y N K K S D V G F	114
GTAAGGTTGATGCTCAAAGAAGACTCTGTTTCTTCTATAACAAGAAATCTGATGTGGGAT	720
L P P F P T L D P E E K C Q A Y E S N R	134
TTCTGCCTCCTTTCCCTACCCTGGATCCCGAAGAGAAATGCCAGGCTTATGAAAGTAACA	780
E S L L N H F L Y E V A R R N P F V F A	154
GAGAATCCCTTTTAAATCACTTTTATATGAAGTTGCCAGAAGGAACCCATTTGTCTTCG	840
P T L L T V A V H F E E V A K S C C E E	174
CCCCTACACTTCTAACTGTTGCTGTTTCATTTTGAGGAGGTGGCCAAATCATGTTGTGAAG	900
Q N K V N C L Q T R A I P V T Q Y L K A	194
AACAAAACAAAGTCAACTGCCTTCAAACAAGGGCAATACCTGTACACAATATTTAAAG	960
F S S Y Q K H V C G A L L K F G T K V V	214
CATTTTCTTCTTATCAAAAACATGTCTGTGGGGCACTTTTGAAATTTGGAACCAAAGTTG	1020
H F I Y I A I L S Q K F P K I E F K E L	234
TACACTTTATATATATTGCGATACTCAGTCAAAAATTCCCCAAGATTGAATTTAAGGAGC	1080
I S L V E D V S S N Y D G C C E G D V V	254
TTATTTCTCTTGTAGAAGATGTTTCTTCCAACATATGATGGATGCTGTGAAGGGGATGTTG	1140

2/12

## FIG. 1B

Q C I R D T S K V M N H I C S K Q D S I 274  
TGCAGTGCATCCGTGACACGAGCAAGGTTATGAACCATATTTGTTCAAAACAAGATTCTA 1200

S S K I K E C C E K K I P E R G Q C I I 294  
TCTCCAGCAAAATCAAAGAGTGCTGTGAAAAGAAAATACCAGAGCGCGGCCAGTGCATAA 1260

N S N K D D R P K D L S L R E G K F T D 314  
TTAACTCAAACAAAGATGATAGACCAAAGGATTTATCTCTAAGAGAAGGAAAATTTACTG 1320

S E N V C Q E R D A D P D T F F A K F T 334  
ACAGTGAAAATGTGTGTCAAGAACGAGATGCTGACCCAGACACCTTCTTTGCGAAGTTTA 1380

F E Y S R R H P D L S I P E L L R I V Q 354  
CTTTTGAATACTCAAGGAGACATCCAGACCTGTCTATAACCAGAGCTTTTAAGAATTGTTTC 1440

I Y K D L L R N C C N T E N P P G C Y R 374  
AAATATACAAAGATCTCCTGAGAAATTGCTGCAACACAGAAAACCCTCCAGGTTGTTACC 1500

Y A E D K F N E T T E K S L K M V Q Q E 394  
GTTACGCGGAAGACAAATTCAATGAGACAACCTGAGAAAAGCCTCAAGATGGTACAACAAG 1560

C K H F Q N L G K D G L K Y H Y L I R L 414  
AATGTAAACATTTCCAGAATTTGGGGAAGGATGGTTTGAAATACCATTACCTCATCAGGC 1620

T K I A P Q L S T E E L V S L G E K M V 434  
TCACGAAGATAGCTCCCCAACTCTCCACTGAAGAACTGGTGTCTCTTGCGAGAAAATGG 1680

T A F T T C C T L S E E F A C V D N L A 454  
TGACAGCTTTCACTACTTGCTGTACGCTAAGTGAAGAGTTTGCCCTGTGTTGATAATTTGG 1740

D L V F G E L C G V N E N R T I N P A V 474  
CAGATTTAGTTTTTGGAGAGTTATGTGGAGTAAATGAAAATCGAACTATCAACCCTGCTG 1800

D H C C K T N F A F R R P C F E S L K A 494  
TGGACCACTGCTGTAAACAAACTTTGCCTTCAGAAGGCCCTGCTTTGAGAGTTTGAAAG 1860

D K T Y V P P P F S Q D L F T F H A D M 514  
CTGATAAAACATATGTGCCTCCACCTTTCTCTCAAGATTTATTTACCTTTACGCAGACA 1920

C Q S Q N E E L Q R K T D R F L V N L V 534  
TGTGTCAATCTCAGAATGAGGAGCTTCAGAGGAAGACAGACAGGTTTCTTGTCAACTTAG 1980

K L K H E L T D E E L Q S L F T N F A N 554  
TGAAGCTGAAGCATGAACTCACAGATGAAGAGCTGCAGTCTTTGTTTACAAATTTGCGAA 2040

V V D K C C K A E S P E V C F N E E S P 574  
ATGTAGTGGATAAGTGCTGCAAAGCAGAGAGTCCTGAAGTCTGCTTTAATGAAGAGAGTC 2100

3/12

## FIG. 1C

K I G N	578
CAAAAATTGGCAACTGAAGCCAGCTGCTGGAGATATGTAAAGAAAAAAGCACCAAAGGGA	2160
AGGCTTCCTATCTGTGTGGTGATGAATCGCATTTCCTGAGAACAAAATAAAAGGATTTT	2220
CTGTAACTGTCACCTGAAATAATACATTGCAGCAAGCAATAAACACAACATTTTGTAAG	2280
TTAAAAA	2287

1

L

X

X

H

X

X

R

D

Y

E

X

X

X

X

F

X

X

L

G

E

E

N

F

X

X

L

X

L

I

A

F

A

Q

Y

V

Q

E

A

T

F

E

E

V

S

K

L

V

K

D

V

V

E

X

X

E

X

C

X

Consensus

Majority

11

L

P

T

Q

P

R

D

I

E

N

F

N

S

T

Q

K

E

I

-

-

E

D

N

I

E

Y

I

T

I

I

A

F

A

Q

Y

V

Q

E

A

T

F

E

E

M

E

K

L

V

K

D

M

V

E

Y

K

D

R

C

M

AFM

111

-

T

L

H

R

N

E

Y

G

I

A

S

I

L

D

S

Y

Q

C

T

A

E

I

S

L

A

D

L

I

I

E

F

A

Q

Y

V

Q

E

A

T

Y

K

E

V

S

K

H

V

K

D

A

L

T

A

I

E

K

P

T

AFP

111

-

D

A

H

K

S

E

-

-

V

A

H

R

F

L

G

E

E

N

F

K

A

L

V

L

I

I

E

F

A

Q

Y

V

Q

E

A

T

F

E

D

H

V

K

L

V

N

E

V

T

E

F

A

K

T

C

V

ALB

111

L

E

-

R

G

R

D

Y

E

K

N

K

V

C

K

E

F

S

H

L

G

R

E

D

F

T

S

L

V

L

Y

S

R

K

F

P

S

G

T

F

E

Q

V

S

Q

L

V

K

E

V

S

L

T

E

A

C

C

VDB

117

A

D

E

X

X

P

X

C

K

X

X

X

A

L

X

E

K

L

C

X

X

X

L

X

E

K

H

G

X

X

A

D

C

C

S

K

X

E

X

E

R

X

L

C

F

L

A

H

K

K

X

X

P

X

X

L

P

Consensus

Majority

160

A

D

E

X

X

P

X

C

K

X

X

X

A

L

X

E

K

L

C

X

X

X

L

X

E

K

H

G

X

X

A

D

C

C

S

K

X

E

X

E

R

X

L

C

F

L

A

H

K

K

X

X

P

X

X

L

P

AFM

155

A

D

E

S

A

E

N

C

D

K

S

L

H

T

L

F

G

D

K

L

C

T

V

A

T

L

R

E

T

Y

G

E

M

A

D

C

C

A

K

Q

S

E

P

E

R

N

H

E

C

F

L

A

H

K

K

D

P

P

AFP

160

A

E

G

A

D

P

D

C

Y

D

T

R

T

S

A

L

S

A

K

S

C

E

S

N

S

P

P

V

H

P

G

T

A

E

C

C

T

K

E

G

L

E

R

K

L

C

-

M

A

A

L

K

H

Q

R

P

-

N

L

P

VDB

1117

X

F

V

X

P

E

P

D

E

X

C

E

A

F

E

X

D

R

E

T

F

L

N

K

F

L

E

I

A

R

R

H

P

F

X

Y

A

P

T

L

L

X

X

A

K

R

Y

X

X

X

X

S

C

C

X

A

A

N

Consensus

Majority

1119

P

F

P

T

L

D

P

E

E

K

C

Q

A

Y

E

S

N

R

E

S

L

I

N

H

F

I

Y

E

V

A

R

R

N

P

F

E

V

E

A

P

T

L

L

T

V

A

V

H

F

E

E

V

A

R

S

C

C

K

E

E

Q

N

AFM

1114

L

E

Q

V

P

E

P

V

T

S

C

E

A

Y

E

D

R

E

T

F

M

N

K

F

I

Y

E

I

A

R

R

H

P

E

L

Y

A

P

T

I

L

L

M

A

A

R

Y

D

K

I

I

P

S

C

C

K

E

A

N

AFP

1119

R

L

V

R

P

F

V

D

V

M

C

T

A

F

H

D

N

E

T

F

L

K

K

Y

L

Y

E

I

A

R

R

H

P

Y

F

Y

A

P

E

L

L

F

F

A

K

R

Y

K

A

A

F

T

E

C

C

Q

A

A

D

ALB

1119

T

Y

V

E

P

T

N

D

E

I

C

E

A

F

R

K

D

P

K

E

A

N

Q

E

M

W

E

S

T

N

Y

G

Q

A

P

L

S

L

L

V

S

Y

T

K

S

Y

L

S

M

V

G

S

C

C

T

S

A

S

VDB

117

C

X

C

L

L

T

K

A

X

X

V

T

X

E

L

K

A

S

S

X

L

O

X

H

V

C

A

X

L

X

K

F

G

T

K

X

F

A

I

X

V

A

K

L

S

Q

K

F

P

K

A

E

F

X

E

V

X

K

Consensus

Majority

1179

K

V

N

C

L

Q

T

R

A

I

P

V

T

Q

Y

L

K

A

F

S

S

Y

Q

K

H

V

C

G

A

L

L

K

F

G

T

K

V

H

F

I

Y

I

A

I

L

S

Q

K

F

P

K

I

E

F

K

E

L

S

AFM

1174

A

V

E

C

F

O

T

K

A

A

T

V

T

R

E

L

R

E

S

S

L

L

N

O

H

A

C

A

V

M

K

N

F

G

T

T

F

Q

A

I

T

V

T

K

L

S

Q

K

F

P

K

V

N

F

T

E

I

Q

K

AFP

1179

K

A

A

C

L

L

P

K

L

D

E

L

R

D

E

G

K

A

S

S

A

K

O

R

L

K

C

A

S

L

Q

K

F

G

E

R

A

E

K

A

V

A

R

L

S

Q

R

F

P

K

A

E

E

A

E

V

S

K

ALB

1179

P

T

V

C

F

L

K

E

R

L

Q

L

K

H

-

L

S

L

L

T

T

L

S

N

R

V

C

S

Q

Y

A

A

Y

G

E

R

K

S

R

L

S

N

L

I

K

L

A

O

R

V

P

T

A

D

L

E

D

V

L

P

VDB

117

L

D

V

T

X

V

H

X

X

C

E

G

D

V

L

D

C

X

-

X

D

X

X

K

X

M

X

Y

I

C

S

N

O

D

S

I

S

S

K

I

K

E

C

C

-

E

K

T

X

L

E

R

G

Q

C

I

Consensus

Majority

1237

L

V

E

D

V

T

X

V

H

X

X

C

E

G

D

V

L

D

C

X

-

X

D

X

X

K

X

M

X

Y

I

C

S

N

O

D

S

I

S

S

K

I

K

E

C

C

-

E

K

T

X

L

E

R

G

Q

C

I

AFM

1239

L

V

E

D

V

S

S

N

Y

D

G

C

C

E

G

D

V

V

Q

C

I

-

R

D

T

S

K

V

M

N

H

I

C

S

K

Q

D

S

I

S

S

K

I

K

E

C

C

-

E

K

K

I

P

E

R

G

Q

C

I

AFP

1234

L

V

L

D

V

A

H

V

H

E

H

C

C

R

G

D

V

L

D

C

L

-

Q

D

G

E

K

I

M

S

Y

I

C

S

Q

D

T

L

S

N

K

I

T

E

C

C

-

K

L

T

T

L

E

R

G

Q

C

I

ALB

1238

L

V

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

I

A

ALB

1238

L

A

E

D

I

T

N

I

L

S

K

C

C

E

S

A

S

E

D

C

A

-

D

D

R

A

D

L

A

K

Y

I

C

E

N

O

D

S

I

S

S

K

L

K

E

C

C

-

E

K

P

L

L

E

K

S

H

C

5/12

FIG. 2A-2

295	NSNKKIDDRPKDLSLRGKFTDSENVCQIERDAADPDTFEAKFTFEYSRRHHPDLSIPELLRIVQ	350	360	AFM
297	HAENDEKPEGLS	352	362	AFP
292	EVENDEMPADLP	354	364	ALB
298	FMPAAQLPE-LP	356	366	VDB
300	Consensus	358	368	Majority
302	XXXXLLEKCCXTEDPXECCYXXKXEDLKKXXEESQXXLXQQXCXLFQXGLGEYXXLQNALLLVRL	360	370	
304	370	370	380	
306	IYKDLLRNCCINTENP	382	392	AFM
308	GYQELLEKCFQNTENP	384	394	AFP
310	357	396	406	ALB
312	352	408	418	VDB
314	351	410	420	
316	Consensus	412	422	Majority
318	TKKAPQLSTXELVXLXRRKMXXXASTCCXLSLSEXX-XACXEXXADXXVXXGXLCXXHXXTPVNP	414	424	
320	430	430	440	
322	TKRIA PQLSTXELVSLGEMVTAFTTCCCTLSE--FACVNDNLADLVFGELCCGVNENRTIINP	432	442	AFM
324	415	444	454	AFP
326	417	456	466	ALB
328	412	468	478	VDB
330	411	470	480	
332	Consensus	472	482	Majority
334	XVXXCCTXSXANRRPCFSSSLXVDETYVPPXFSXDXTFFHADXCQSQXXELOXXKKQXFLVN	474	484	
336	490	490	500	
338	473	502	512	AFM
340	477	514	524	AFP
342	472	526	536	ALB
344	442	538	548	VDB
346	Consensus	540	550	Majority
348	LVRKXRPXXXTDEELKAVXXDFAXXVEXKCCCKAXXXEVCFAEEGXKLLXXXAAALGX	542	552	
350	550	550	560	
352	533	562	572	AFM
354	537	574	584	AFP
356	532	586	596	ALB
358	450	598	608	VDB

## FIG. 2B

		<u>% Similarity</u>			
		<u>AFM</u>	<u>AFP</u>	<u>ALB</u>	<u>VDB</u>
<u>% Identity</u>	<u>AFM</u>		60.4	54.8	41.2
	<u>AFP</u>	39.6		59.1	40.9
	<u>ALB</u>	35.6	40.2		45.8
	<u>VDB</u>	20.7	21.2	24.4	



7/12

FIG. 3

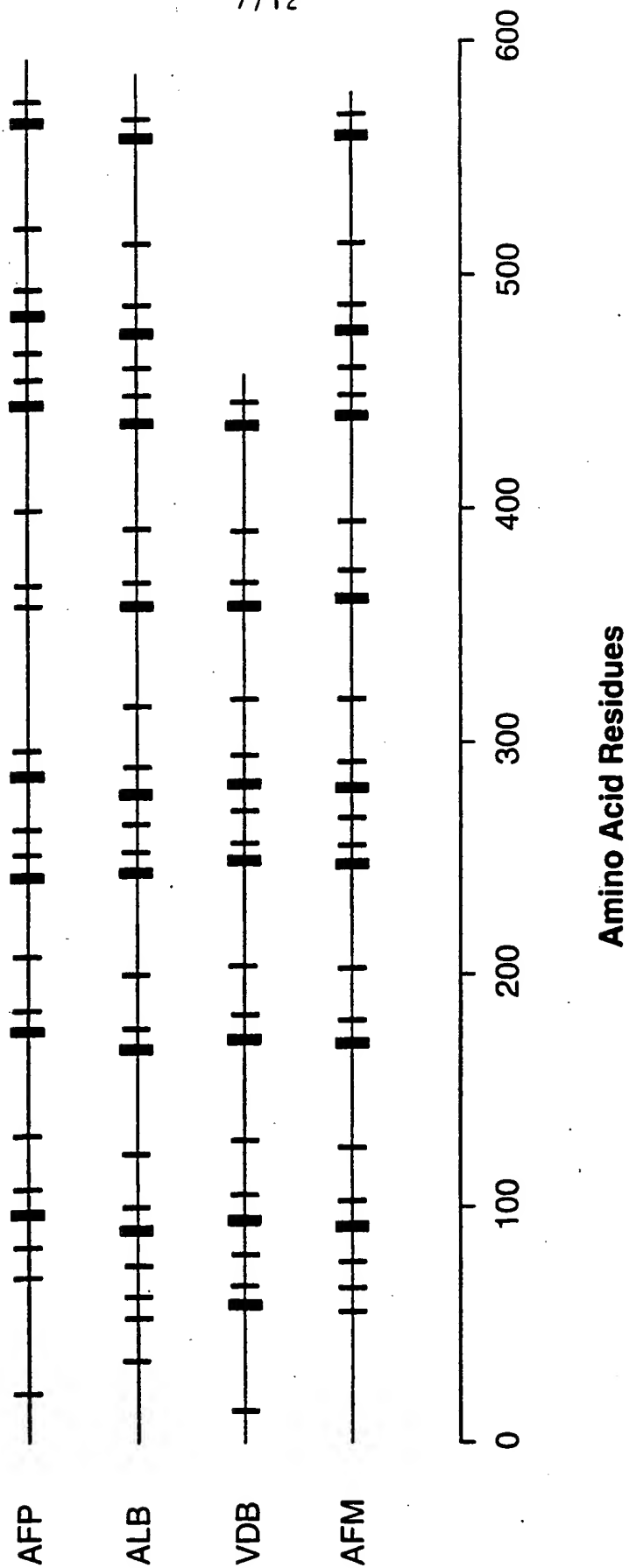
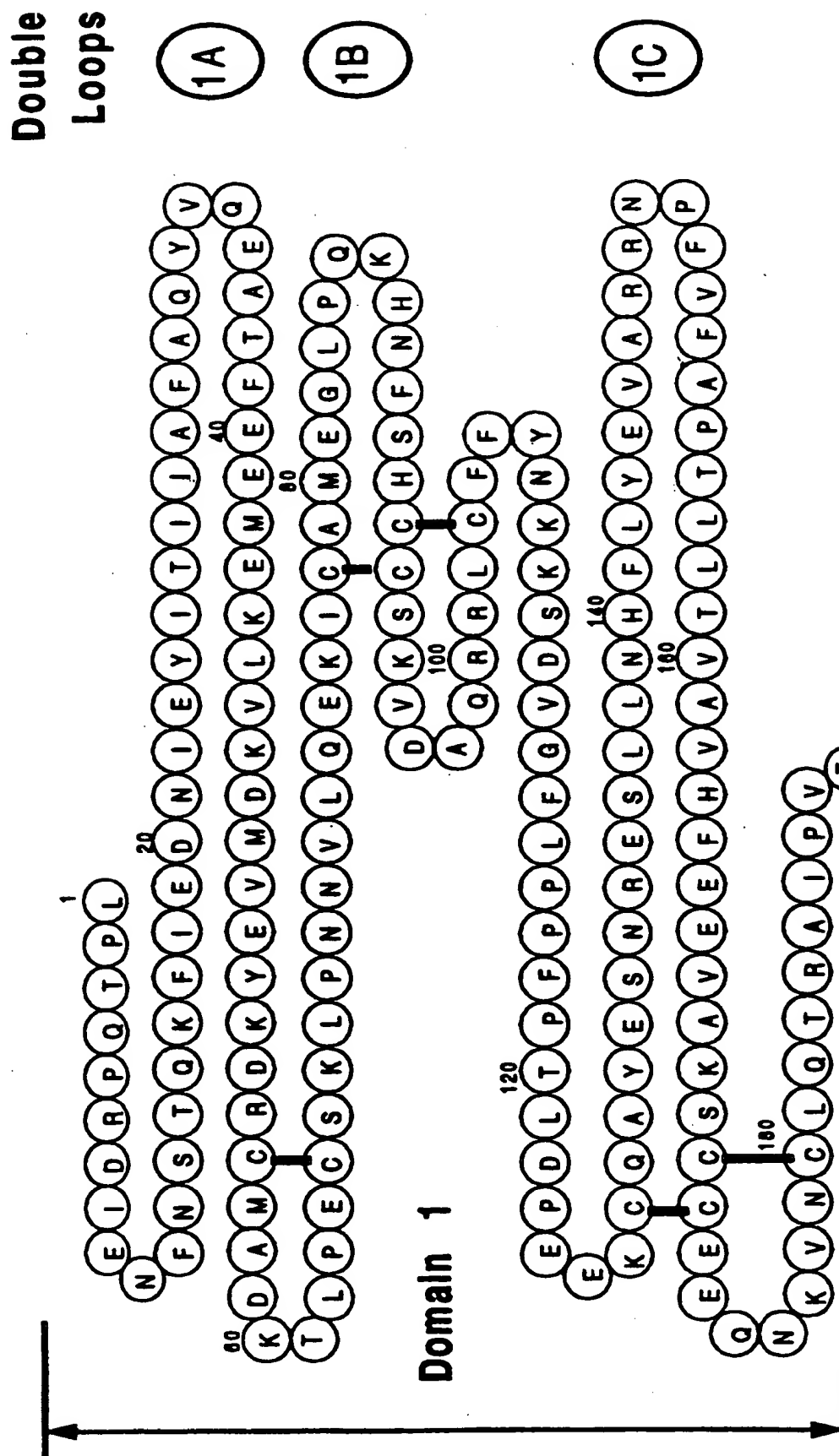


FIG. 4A



9/12

FIG. 4B

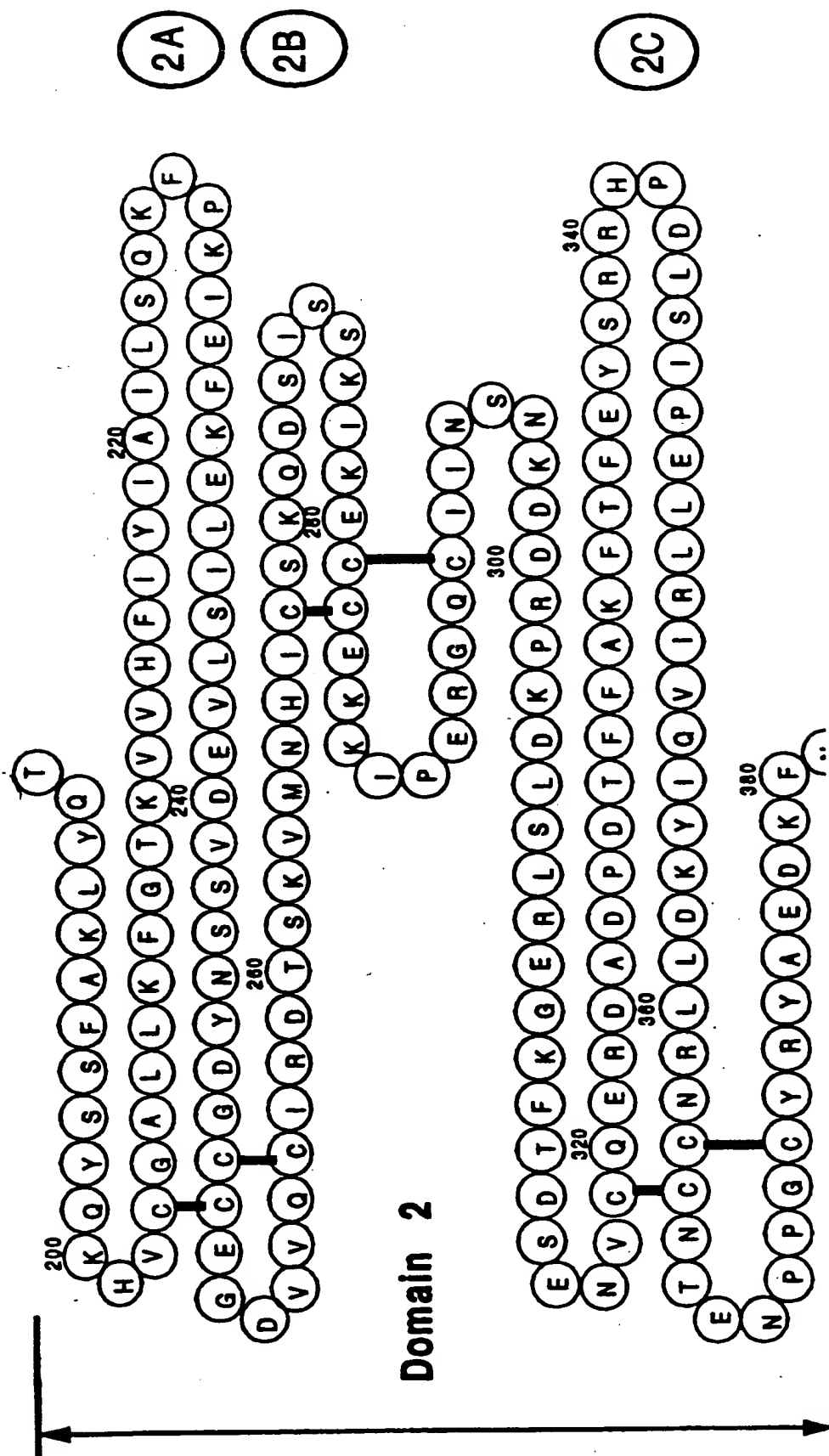
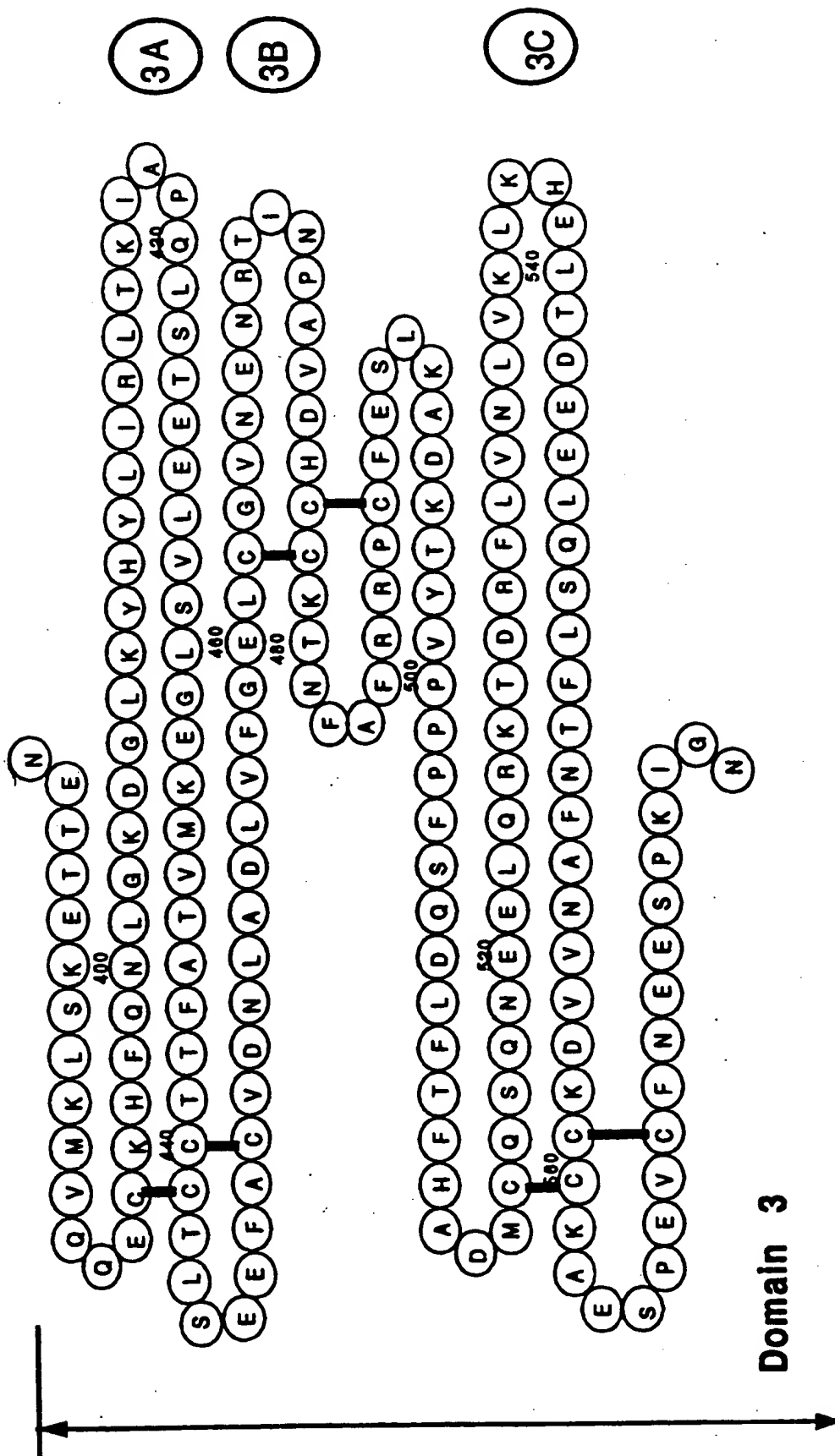
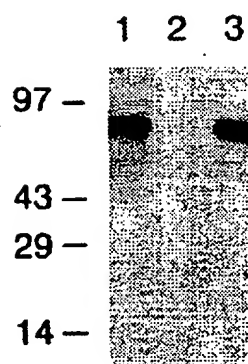


FIG. 4C



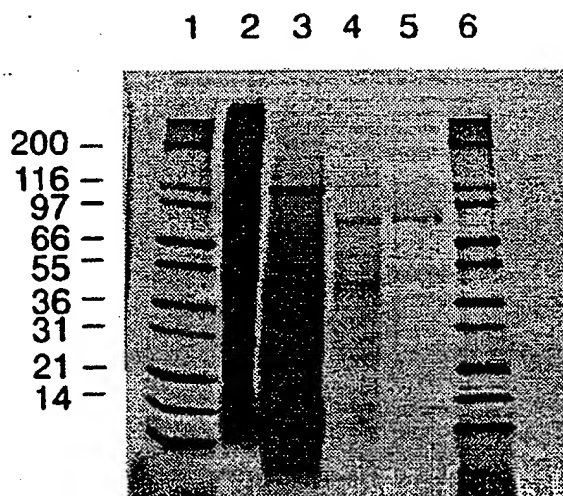
11/12

FIG. 5



12/12

FIG. 6



## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/62 C07K14/47 C07K16/18 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 269, no. 8, 25 February 1994 pages 5481-5484, BÉLANGER ET AL 'NEW ALBUMIN GENE 3' ADJACENT TO THE ALPHA1-FETOPROTEIN LOCUS' see the whole document ---	1-27
P,X	THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 269, no. 27, 8 July 1994 pages 18149-18154, LICHENSTEIN ET AL 'AFAMIN IS A NEW MEMBER OF THE ALBUMIN, ALPHA-FETOPROTEIN, AND VITAMIN D-BINDING PROTEIN GENE FAMILY' see the whole document --- -/--	1-27



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*I\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*A\* document member of the same patent family

Date of the actual completion of the international search

2 August 1995

Date of mailing of the international search report

1 0. 08. 95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Sitch, W

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	<p>WO,A,95 11308 (AMGEN INC) 27 April 1995</p> <p>see page 6, line 20 - page 35</p> <p>see SEQ ID NOS: 23 and 24</p> <p>see page 41 - page 43; example 4</p> <p>-----</p>	<p>1-14, 16, 27</p>



